University College London

A Computational Investigation of the Interaction of the Collagen Molecule with Hydroxyapatite

Thesis submitted for the degree of Doctor of Philosophy (PhD) by

Neyvis Almora Barrios

Supervisors:

Prof. Nora H. de Leeuw

University College London

Department of Chemistry

January 2010

Declaration:

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Neyvis Almora Barrios

February 2010

Abstract

This thesis presents the results of computer simulation studies of the interaction of the predominant molecules in the collagen protein with the hydroxyapatite mineral. Using a combination of computational techniques, quantum-mechanical methods based on the density functional theory (DFT) and molecular dynamics simulations based on interatomic potentials, we have investigated the interface between the collagen protein and the apatite mineral.

First we have employed electronic structure techniques (DFT) to study a range of different binding modes of the amino acids glycine, proline and hydroxyproline, which are major constituents of the collagen I protein, at two important hydroxyapatite surfaces, (0001) and $(01\bar{1}0)$. We have performed full geometry optimizations of the hydroxyapatite surface with adsorbed amino acid molecules to obtain the optimum substrate/adsorbate structures and interaction energies. We have also used DFT to investigate the binding of a series of representative peptides containing hydrophobic side groups (proline), uncharged polar side groups (glycine and hydroxyproline), and charged polar side groups (lysine and hydroxylysine) to the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces. This selection of adsorbates has given us the opportunity to study separately the interactions of the carboxylic acid and amine functional groups, as well as the effect of hydroxylation and the charges of the side group, on the strength of interaction with the surfaces.

We have also investigated the same systems in an aqueous environment using classical molecular dynamics simulation, where we have calculated the energies

and geometries of adsorption of the peptide at the surfaces of hydroxyapatite in competition with pre-adsorbed water. Finally, we have studied the onset of nucleation of the hydroxyapatite mineral at an entire collagen molecule in aqueous solution.

Table of Contents

Acknowledgments

First I want to thank my supervisor, Prof. Nora de Leeuw, for her guidance and support during these three years.

I am very grateful for the UCL Dorothy Hodgkin Overseas Research Studentship, which generously provided the funds for the execution of this project and for my university fees and maintenance.

Special acknowledgments should also go to Dr. Ricardo Grau-Crespo for his invaluable help, guidance and all his patience during these years. I would not have been able to complete this work without his support. I have also received valuable input from Dr. Devis di Tommaso, Dr. Rabdel Ruiz-Salvador, Dr. Iman Saadoune, Dr. Zhimei Du, Will Richards, and all members of the lunch group.

I am very grateful to my mum, dad, grandparents and sisters for their love, support and encouragement. Finally, I would like to thank my son Alejandro to whom I dedicate this work.

List of Publications

The work described in this thesis has been published in the following papers:

N. Almora-Barrios, K. Austen, N. de Leeuw, "A Density Functional Theory study of the binding of glycine, proline and hydroxyproline to the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces", *Langmuir* (2009), 25, 5018-5025

N. de Leeuw, **N. Almora-Barrios,** "Density Functional Theory Calculations and Molecular Dynamics Simulations of the Interaction of Bio-molecules with Hydroxyapatite Surfaces in an Aqueous Environment", *Materials Research Society* (2009) DOI: 10.1557/PROC-1131-Y01-06

N. Almora-Barrios, N. H. de Leeuw, "Computer simulations of the interaction of bio-molecules with hydroxyapatite surfaces in aqueous environment", in *Geochim. Cosmochin. Acta* (2009) A31, 73.

N. Almora-Barrios, N. de Leeuw, "Modelling the interaction of a Hyp-Pro-Gly peptide with hydroxyapatite surfaces in aqueous environment", *CrystEngComm*, (2010), DOI: 10.1039/b917179g.

N. Almora-Barrios, N. de Leeuw, "A Density Functional Theory study of the binding of the collagen peptide to the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces", (submitted) (2010).

Chapter 1 Biomaterials: collagen and apatite

A number of materials, ranging from metals to ceramics, have been used for tissue replacement in dental and orthopaedic applications, where they are often designed to resemble as closely as possible the natural tissue. Understanding at the atomic level the interaction of a protein with an inorganic surface is very important in the development of these biomaterials, especially for biocompatible medical implants.

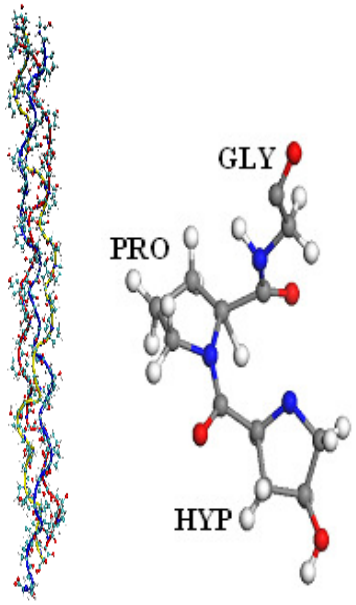
1.2 Hierarchical structure of bone

Natural bone material is a highly hierarchical protein-mineral composite, containing nano-sized mineral platelets (essentially carbonated hydroxyapatite), proteins (predominantly collagen) and water (Fratzl *et al.*, 2004). Although the mineral phase is stiff and brittle, while the protein is much softer and tougher, the composite combines the optimal properties, stiffness and toughness, of both components. This unusual combination of material properties provides both rigidity and resistance against fracture.

1.1.1 Collagen molecule

The protein in bone is predominantly type I collagen, whose structure is complex and is self-assembled in a multi-step process (Stevens & George, 2005). The primary structure of the collagen contain the complete sequence of amino acids along each of three polypeptide chains that are coiled into a left-handed helix, where hydrogen bonding plays a structure directing role. The three chains are then

wrapped around each other into a right-handed super-helix so that the final structure is a rope-like rod. The presence of glycine as the third amino acid in the repeating GLY-X-Y- sequence of each chain is essential, because a larger amino acid will not fit in the centre of the triple helix, where the three chains come together. Proline is frequently in the X-position of the -GLY-X-Y- sequence and 4-hydroxyproline is frequently in the Y-position (figure 1.1).



(a) **(b)**

Figure 1.1 (a) Structure of the triple helix of three polypeptide collagen chains with a highly repetitive amino acid sequence, **(b)** $[GLY\text{-}PRO\text{-}HYP]_n$ (C = grey, O = red, N = blue, H = white).

The X and Y amino acids limit rotation of the polypeptide chains, and the triple helix is further stabilised by hydrogen bonds and water bridges, many of which require the presence of 4-hydroxyproline. The conformation of the triple helix places the side chains of the amino acids in the X- and Y- position on the surface of the molecule (Prockop & Kivirikko, 1995), where they can interact with

neighbouring chains or other molecules. The structure of the bifunctional interchain cross-link is the relatively complex condensation product of a reaction involving lysine and hydroxylysine residues; this reaction continues as the organism matures, thereby rendering the collagens of older animals more difficult to extract from tissues.

The tertiary structure of the collagen refers to the global configuration of the polypeptide chains. It represents the pattern in which the secondary structure is packed within the complete macro-molecule and it constitutes the structural unit that can exist as a physicochemically stable entity in solution, namely, the triple helical collagen molecule.

The biological assembly of the collagen fibre requires many steps (figure 1.2). The collagen molecule is synthesized in form of a precursor called procollagen in fibroblast cell. The subunits of the procollagen contain extension peptides at their amino and carboxyl termini. Next, procollagen chains are aggregated in triple-stranded bundles. Extension peptides are necessary for the proper alignment of chains but these are removed after procollagen bundles are secreted into the extracellular matrix. Microfibrils form when collagen molecules aggregate in the extracellular matrix near the surface of the fibroblasts. Finally, these collagen molecules are packed in a quasi-hexagonal lattice at an interchain distance of about 1.3 nm, which is a quarter of its length from its neighbouring molecules in

the bundle so that there is completes overlapping only at intervals of five molecules.

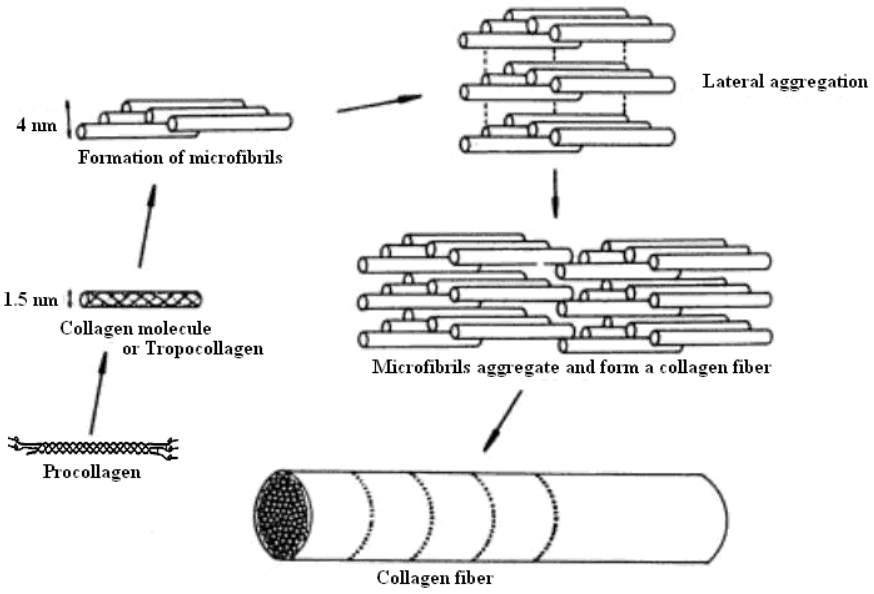


Figure 1.2 Schematic for the structure of collagen fiber.

In microfibril aggregates, there is a distance between the C-terminus and N-terminus of consecutive collagen molecule of about 32 nm. In natural bones, this space, which may be important in bone formation, is filled with a calcium phosphate, hydroxyapatite.

1.1.2 Structure of Apatites

Apatites $Ca_{10}(PO_4)_6(F, Cl, OH)_2$ are a group of phosphate minerals, which are becoming increasingly important. Apatite is one of few minerals that are produced and used by biomaterials. Also it is occasionally found to contain significant amounts of rare earth elements and can be used as alternative solid electrolyte

materials with potential use in fuel cell ceramic membranes (Nakayama *et al.*, 1995, Tolchard *et al.*, 2007, Jones *et al.*, 2008).

In bone, as already mentioned, the inorganic mineral phase consists of small crystallites of carbonate apatite, which is essentially the calcium phosphate mineral hydroxyapatite $Ca_{10}(PO_4)_6(OH)_2$ (HA), where carbonate groups have replaced a greater or lesser fraction of phosphates and/or hydroxy groups in the structure (figure 1.3) (Narasaraju & Phebe, 1996).

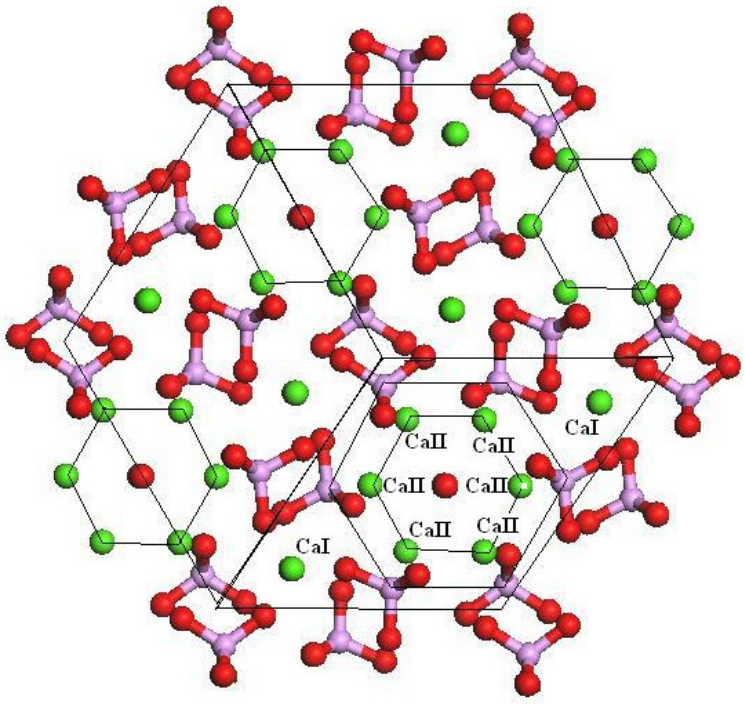


Figure 1.3 Hydroxyapatite structure viewed onto the (0001) surface (Ca = green, O = red, P = purple, H = white).

Natural hydroxyapatite has a hexagonal crystal structure with space group $P6_3/m$ (figure 1.4). The *a* and *c* parameters for hydroxyapatite are 9.423 Å and 6.883 Å,

respectively (Posner *et al.*, 1958). Four calcium ions (Ca I) in the unit cell are usually labeled as columnar calcium because they form single atomic columns perpendicular to the basal plane. Six of the ten calcium ions (Ca II) in the unit cell form hexagonal channels along the *c*-direction of the structure, where are at the vertices of triangles around each hydroxyl group (OH). Four OH positions per unit cell are available in each channel, alternately pointing in opposite directions, leading to a 0.5 partial occupancy of each position. The thermodynamically preferred configuration consists of columns of aligned OH groups within each channel, although neighbouring columns may be aligned in opposite directions (de Leeuw, 2001). This structure is similar to the monoclinic phase of synthetic HA, which is, however, not found in nature.

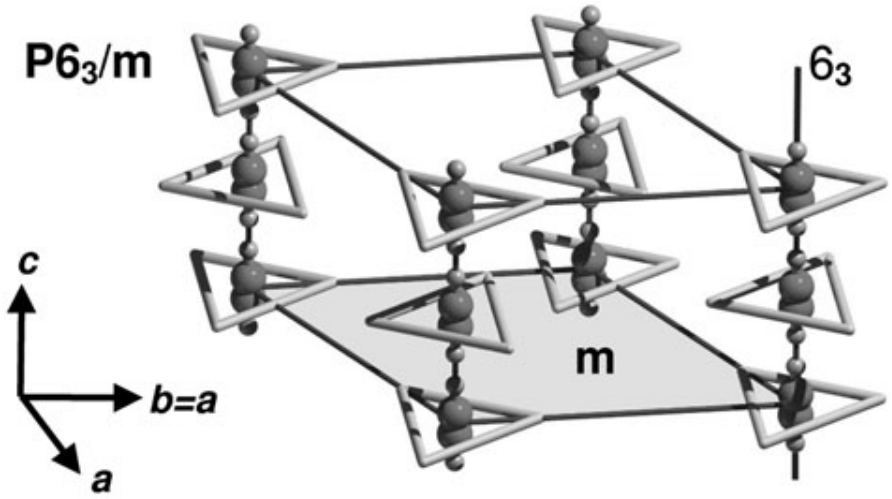


Figure 1.4 Experimental hexagonal $P6_3/m$ hydroxyapatite structure. P, O and Ca atoms not relevant for the scheme have been omitted for clarity. Ca II atoms are at the vertices of triangles around each hydroxyl group.

Defects and impurities in biological apatites may be identified as substitutional or as discrete, extraneous crystalline phases. Ions that may be incorporated into the apatite structures, either intentionally or unintentionally, include carbonate ions (Barralet *et al.*, 1998, Zapantal.R, 1965) (substituting for hydroxy or phosphate groups), fluoride ions (Jha *et al.*, 1997) (substituting for hydroxy groups), silicon, or silicate ions (Gibson *et al.*, 1999) (substituting for phosphorus or phosphate groups) and magnesium and strontium ions substituting for calcium (Bertinetti *et al.*, 2009, Drouet *et al.*, 2008).

1.2.3 Apatite/collagen composite

The mineral phase in bone is deposited within the collagen fibrils as minute needles or platelets of dimensions 10-60 nm in length and 2-6 nm in width, such that their long axes line up with the collagen matrix. A sketch of the likely arrangement of the mineral particles with respect to the collagen molecules is show in Figure 1.5. These mineralized fibrils are then arranged in parallel arrays, and the arrays are further organized into three-dimensional structures to form the bone tissue (Vanderby, 2003). The precise structure and orientation of the mineralised collagen fibril in the bone largely depends on its function; regions experiencing tensile loading are inclined to have more longitudinally oriented fibers, while regions under compressive loading typically contain more transverse fibers (Vanderby, 2003, Martin *et al.*, 1998).

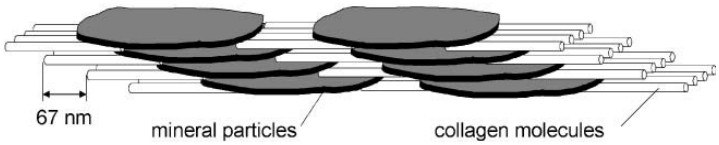


Figure 1.5 Sketch of the arrangement of mineral particle in collagen fibrils. (from reference (Fratzl *et al.*, 2004))

In summary, natural bone tissues are built up of many distinct levels with highly specific interactions between these levels as show in Figure 1.6, and they have the architecture to accommodate a complex spectrum of property requirements (Sarikaya & Aksay, 1995). However, the detailed mechanisms of interaction between collagen and hydroxyapatite are still not clear, which has motivated a number of both experimental and theoretical studies (Elliott *et al.*, 1973, Mann, 2001, Simon *et al.*, 2006, Tlatlik *et al.*, 2006). The understanding of such interactions is crucial in the development of topical research fields, such as biomineralization, nanomaterials, and biochemistry.

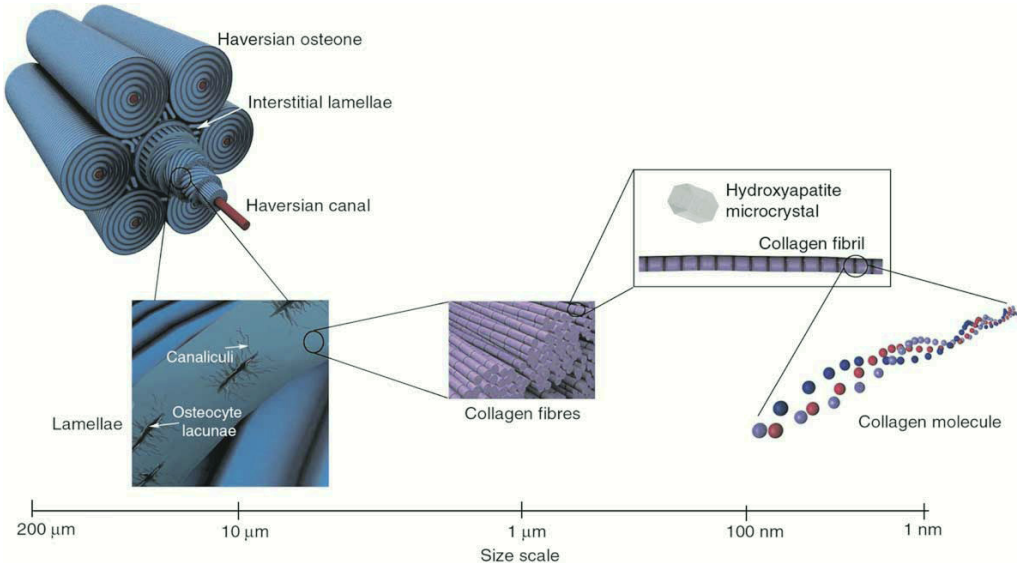


Figure 1.6 Human bone is a hierarchical composite of different constituents which are employed in the construction of bone tissue of two varieties, namely woven and lamellar bone (from reference (Tirrell *et al.*, 2002))

1.2 Biomaterials

Due to the optimal properties of the natural bone tissue, a number of composite materials have been investigated as possible bio-active implant materials for bio-medical applications, for example polymer/apatite and polymer/bio-glass composites (Boccaccini *et al.*, 2005). Bio-active implant materials such as Bioglass®, first discovered by Hench and co-workers, have the tendency to dissolve in the body and be resorbed into hard tissue (Sanders *et al.*, 1972). As such, these materials do not work themselves loose, as the older, bio-inert implants eventually do, and hence they need not be replaced regularly in order to eliminate the need for invasive repeat surgery. Since the discovery of the bio-compatible properties of Bioglass® have revolutionized the field of bio-materials research, most investigations have concentrated on designing bio-active implant materials for hard tissue replacement, which resemble as closely as possible the structure and properties of natural bone, including its composite nature.

Although Poly (D,L lactic acid), polyethylene and poly-ethylene glycol foams are widely used in composite materials for clinical and pharmaceutical applications (Quirk *et al.*, 2004, Barry *et al.*, 2005), collagen itself has also been investigated as a bio-material candidate (Qian & Bhatnagar, 1996, Fields *et al.*, 1998, Johnson *et al.*, 2000, Wang *et al.*, 2005). In this case the triple helices are limited to the length of a synthetic peptide (~10 nm), which is much shorter than the natural collagen (~300 nm) (Kotch & Raines, 2006). Collagen-based biomedical devices have been developed in a variety of physical forms, including nanoparticles,

fibres, films, hydrogels and pellets, where they have been used as drug carriers, wound dressings, skin replacements and bone substitutes (Lee *et al.*, 2001). However, in the most recent bio-materials applications, the collagen matrix is usually combined with a reinforcing phase such as hydroxyapatite, which should improve both the mechanical properties and the bio-activity of the material (Eglin *et al.*, 2006).

1.3 Experimental Studies of the composite organic-apatite materials

Apatite-collagen composites are biodegradable and good matrices for bone cell attachment and proliferation as well as new bone formation (Tirrell *et al.*, 2002). Experimental studies have demonstrated that the growth process of apatite is controlled by the interaction between the organic matrix and the apatite crystal (Du *et al.*, 2005, Kirkham *et al.*, 2002, Minton, 1999, Smith & Nanci, 1996, Wen *et al.*, 2001), which determines the eventual morphology of the HA platelets in the bone.

Based on electron microscopy and computer simulation studies, recent investigations of biomimetic apatite–gelatine composites have revealed a series of insights into the growth mechanisms at the mesoscopic scale (Simon *et al.*, 2005). On the nm scale, the morphogenesis of the composite is controlled by the arrangement of the biomolecules, the polar nature of which gives rise to an intrinsic dipole field (Simon *et al.*, 2006). A prismatic seed is initially formed; with increasing size the composite experiences increasing electrostatic

interactions which induce branching patterns, and finally lead to the development of spherical morphologies (Kniep & Simon, 2008). Electron microscopy has proved to be a powerful tool for identifying nanoscale patterns of apatite–collagen composites (Simon *et al.*, 2005, Kniep & Simon, 2008).

Crystallographic methods may only provide insights in the inorganic, crystalline component of biomaterials. Various methods have been investigated to produce synthetic hydroxyapatite. The commercial routes are based on aqueous precipitation or conversion from other calcium compounds. On the other hand, a large number of investigations examining bound proteins have used techniques that give limited information. These include fluorescence (Chen *et al.*, 1999, Prokopowicz *et al.*, 2005), atomic force microscopy (Marchin & Berrie, 2003, Agnihotri & Siedlecki, 2004, Cacciafesta *et al.*, 2000), mass spectrometry (Xia & Castner, 2003), nuclear magnetic resonance (NMR) (Lundqvist *et al.*, 2005), and antibody or platelet binding (Grunkemeier & Horbett, 1996, Balasubramanian *et al.*, 1999). Most of the analytical tools used give results where generalised information on the adsorbed protein is obtained or conformational changes are implied but these have not been measured or quantified.

1.4 Theoretical studies of the organic-inorganic interface

The organic-inorganic interface is important since it provides the active sites for the biological control of biomaterials. An important but largely unresolved issue is the way in which nature controls the nucleation, growth, and morphology of

inorganic crystallites and the function of biomolecules (such as amino acids and proteins) in these reactions. However, it is not yet possible to study directly, by experiment alone, the molecular mechanisms of fast processes such as the binding sites and adsorption modes of these biomolecules on the HA surfaces. A wide variety of simulation techniques have been developed, applicable to a range of different problems in biomolecular science.

Simulations are a useful tool for modelling and probing the structure and adsorption processes of organic molecules on inorganic crystals at the atomic level (Harding & Duffy, 2006, Cooper & de Leeuw, 2004, de Leeuw & Cooper, 2004, Nair *et al.*, 2006, de Leeuw, 2004b, de Leeuw & Mkhonto, 2003, Mkhonto & de Leeuw, 2002). Useful information, such as the interfacial energy (Duffy & Harding, 2004b), adsorption energy (Harding & Duffy, 2006, Pan *et al.*, 2007), and interaction sites (de Leeuw & Cooper, 2004), can be obtained from molecular dynamics simulations to gain insight into organic-inorganic interfaces. More recently, *ab initio* methods have been applied successfully to simulate the adsorption of protein residues at the inorganic materials (Nair *et al.*, 2006, Almora-Barrios *et al.*, 2009, Rimola *et al.*, 2008, Borisenko *et al.*, 2008), but this method cannot generally be taken beyond a few hundred atoms for a few tens of picoseconds. Alternatively, *ab initio* methods have often been employed to derive reliable interatomic potential models for the less compute-intensive interatomic potential-based techniques, which allow for the simulation of larger and more complex systems, such as the composite materials studied in this thesis.

Previously, empirical force fields have been established for the simulation of apatite crystals (de Leeuw, 2004b, Hauptmann *et al.*, 2003, Meis *et al.*, 2000, Cruz *et al.*, 2005), where the simulations results were shown to be in good agreement with the experimental crystal structure (de Leeuw, 2004b, Hauptmann *et al.*, 2003, Meis *et al.*, 2000, Cruz *et al.*, 2005, Mkhonto & de Leeuw, 2002), IR spectra (Hauptmann *et al.*, 2003), and molar enthalpy (Cruz *et al.*, 2005, Cruz *et al.*, 2006a, b). Because of the importance of water in both biological and mineral systems, a number of the force fields have been developed to deal with the water interactions, such as rigid model TIPnP and SPCE, and the shell-model water potential models fitted by de Leeuw and Parker (de Leeuw & Parker, 1998). The interfacial structure of the apatite/water interface has been investigated by de Leeuw *et al.* (de Leeuw, 2004b, Mkhonto & de Leeuw, 2002, de Leeuw & Rabone, 2007), and studies of the adsorption processes of biomolecules to apatite surfaces have been reported by several authors (Chen *et al.*, 2007, Pan *et al.*, 2007, Capriotti *et al.*, 2007, Shen *et al.*, 2008). These studies have highlighted the rôle of $–COO^-$ and $–NH_3^+$ groups on the interaction energy.

In particular, molecular dynamic simulations appear to be a very promising tool for the study of biomaterials. Detailed structural information such as the binding sites and adsorption modes of collagen biomolecules to the apatite surfaces is still largely deficient. Quantitative knowledge of the collagen/apatite interface is required to understand the nucleation and directed growth of hydroxyapatite at the collagen matrix. This thesis therefore aims to fill this gap and contribute to a

better understanding of the interaction of collagen with hydroxyapatite at the atomic level, by carrying out an extensive computational study of these materials.

1.5 Nucleation and growth

The nucleation and crystal growth process can occur either spontaneously in solution or due to the presence of foreign particles, which can be naturally there or could have been introduced in the solution in order to act as nucleation sites. The first process is called homogeneous nucleation, and the latter, aided by foreign particles, is called heterogeneous nucleation.

In the first step of the nucleation process, molecular clusters aggregate stepwise until a critical cluster is formed, as follows:

$$X_1 + X_1 \underset{}{\overset{\omega_1^+/\omega_2^-}{\rightleftharpoons}} X_2$$
$$X_2 + X_1 \underset{}{\overset{\omega_2^+/\omega_3^-}{\rightleftharpoons}} X_3$$
$$\vdots$$
$$X_n + X_1 \underset{}{\overset{\omega_n^+/\omega_{n+1}^-}{\rightleftharpoons}} X_{n+1}$$
$$\vdots$$

where ω_n^+ is the rate constant of direct reaction, ω_n^- is the rate constant for inverse reaction and X_1 is the chemical symbol of the single building unit (atom, ionic or molecule). Once nuclei are formed, crystal growth can occur to form crystallites. The crystal growth rate depends on the degree of saturation in solution.

1.5.1 Heterogeneous nucleation

The investigation of surface processes is important for the understanding of heterogeneous precipitation. Thermodynamically, the (0001) surface is the most stable in hydroxyapatite, but the $(01\bar{1}0)$ plane is dominant in the morphology of the biological material, due to the growth-directing effect of the collagen matrix.

The success of heterogeneous nucleation methods resulted in studies into the influence of several ions, such as magnesium (Amjad *et al.*, 1984), cadmium (Dalas & Koutsoukos, 1989), strontium (Koutsoukos & Nancollas, 1981), lithium (Koutsoukos & Nancollas, 1986), and zinc (Dalpi *et al.*, 1993) on the precipitation kinetics on different surfaces of hydroxyapatite. These studies showed inhibition of growth in the presence of these ions, which was mostly related to the adsorption of the ion to available surface sites.

The influence of organic molecules upon the precipitation of calcium phosphate has also been of medical interest for a long time. Amino acids, peptides and proteins, in particular, are suspected to influence the (bio-) apatite formation. Recent works have reported that they are effective regulators in processes involving both the nucleation and crystal growth of hydroxyapatite (Pan *et al.*, 2007, Koutsopoulos & Dalas, 2000a, Magne *et al.*, 2001, Chen *et al.*, 1984, Koutsopoulos & Dalas, 2000b, Jack *et al.*, 2007, George & Veis, 2008). High-resolution electron microscopy studies of hydroxyapatite crystals in tooth enamel,

for example, have shown that matrix proteins affect the growth of the $(01\bar{1}0)$ surface only and not of the other hydroxyapatite surfaces (Bres & Hutchison, 2002). However, detailed structural information such as the binding sites and adsorption modes of these biomolecules to the hydroxyapatite surfaces is still largely lacking. Quantitative knowledge of the collagen/apatite interface is required to understand the nucleation and directed growth of hydroxyapatite at the collagen matrix.

1.6 Objectives of the thesis

In view of the importance of the apatite/collagen composite in both natural bone tissue and in synthetic bio-materials, we have employed a range of computer simulation techniques to perform a quantitative study of the interactions at the interface between the collagen protein and the apatite mineral.

The main objectives of this thesis are:

- To provide energetic and structural details of the interactions at the atomic level between the hydroxyapatite surfaces and the adsorbed amino acids or peptides;
- To understand the role of the water in these interactions;
- To gain insight into the role of collagen as a template in apatite nucleation and growth;

This information will be relevant to our understanding of *in vivo* bone growth, will contribute to rationalise the synthesis of artificial bone, and will expand our general knowledge of bio-mineralisation processes.

Chapter 2 Computational Techniques

In this chapter we describe the computational chemistry methodologies, which have been applied in this study. The computational methods can roughly be divided into two techniques:

- Electronic structure calculations, which involve solving the Schrödinger equation for the investigated system to some level of approximation, have the objective of determining all properties from first principles.
- Atomistic simulation based on the description of interatomic forces using analytical functions, the accuracy of which dictates the quality of the simulation result.

2.1 Electronic Structure Methods

According to the laws of quantum mechanics state that the electronic and structural properties of a given material, with M nuclei and N electrons, can, in principle, be calculated by solving Schrödinger's equation. The time independent Schrödinger's equation has the form:

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian, a differential operator representing the total energy of the system, E is the energy of the state, and Ψ is the wavefunction. The latter

depends on all degrees of freedom, i.e. the coordinates of all particles in the system.

The non-relativistic Hamiltonian operator $\hat{H}$ is defined as the sum of the kinetic and potential operators of all the particles of the system (electrons and nuclei):

$$\left[-\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \right] \Psi = E\Psi \qquad (2.2)$$

given here in atomic units (lengths in Bohrs, masses in electron masses, charges in elementary charges, and energies in Hartrees). The indexes *i* and *j* enumerate the electrons and run from 1 to *N*, while the indexes *A* and *B*, running from 1 to *M*, enumerate the nuclei; M_A and Z_A are the mass and the atomic number of the nucleus *A*, and r_{iA}, r_{ij}, R_{AB} are the electron-nucleus, electron-electron and nucleus-nucleus distances, respectively. The Laplacian operators ∇_i^2 and ∇_A^2 involve differentiation with respect to the coordinates of the ith electron and of the Ath nucleus. The first term in equation 2.2 is the operator for the kinetic energy of the electrons; the second term is the kinetic energy of the nuclei; the third term is the coulomb interaction between electrons and protons; the fourth and the fifth the repulsion between electrons and protons, respectively.

The Schrödinger's equation, however, can only be solved exactly for hydrogen atoms and one-electron systems, whereas for systems with two or more electrons approximations are needed. It is usual to invoke the so-called adiabatic or Born-

Oppenheimer approximation (Born & Oppenheimer, 1927): Since there is a large difference in mass between the electrons and nuclei, the electrons respond essentially instantaneously to the motion of the nuclei. This approximation makes it possible to separate the electronic and nuclear coordinates in the many-body wavefunction, reducing the problem to the solution of the dynamics of the electrons in some frozen configuration of the nuclei.

Within this approximation, the second term of equation 2.2, the kinetic energy of the nuclei, is neglected and the last term, the repulsion between the nuclei, is considered to be constant. The remaining terms of 2.2 define the so-called electronic Hamiltonian, describing the motion of the N electrons in the field of the M point charges:

$$\hat{H}_{elec} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{2.3}$$

The solution of Schrödinger's equation involves the electronic Hamiltonian

$$\hat{H}_{elec} \Psi_{elec} = E_{elec} \Psi_{elec} \tag{2.4}$$

Ψ_{elec} is the electronic wavefunction:

$$\Psi_{elec} = \Psi_{elec} \left(\left\{ \overrightarrow{X_i} \right\}; \left\{ \overrightarrow{R_A} \right\} \right) \tag{2.5}$$

and E_{elec} is the electronic energy:

$$E_{elec} = E_{elec}\left(\left\{\overrightarrow{R_A}\right\}\right) \tag{2.6}$$

which defines parametrically on the nucleus coordinates.

From the resolution of the electronic problem (equation 2.4) in a fixed nuclear configuration $\left\{\overrightarrow{R_A}\right\}$, we can obtain the potential generated by the electrons on the nuclei, which allows the calculation of the forces acting on the nuclei. This represents the basis for the methods of geometry optimization and *ab initio* molecular dynamics. The surface defined by $E_{elec} = E_{elec}\left(\left\{\overrightarrow{R_A}\right\}\right)$ is called the Born Oppenheimer surface: at equilibrium the geometry of a material corresponds to a minimum of this surface.

Despite the Born and Oppenheirmer approximation, solving equation (2.4) is still a formidable task, even for simple atoms or molecules. The main problem is related to the presence of the interaction term $\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{1}{r_{ij}}$. The approach is to define for each ith electron of the system an effective potential $\hat{h}_i$ associated with the presence of the other electrons. This allows re-writing the total N-electron Hamiltonian $\hat{H}_{elec}$ as the sum of N single-electron Hamiltonians $\hat{h}_i$:

$$\hat{H}_{elec} = \sum_{i=1}^{N} \hat{h}_i \tag{2.7}$$

There are two common schemes that allow reformulating the multi-electron problem as equation 2.7:

- Hartree-Fock
- Density Functional Theory

2.1.1 Hartree Fock theory

Hartree Fock (HF) theory uses the original multi-electron wavefunction for an atom as the product of one-electron orbitals $\psi_i(\mathbf{x})$ in the following way:

$$\Psi^{HF}(\mathbf{x}_1,...,\mathbf{x}_N) = \psi_i(\mathbf{x}_1)\psi_j(\mathbf{x}_2)...\psi_k(\mathbf{x}_N) \tag{2.8}$$

The wavefunction of the system is also shared in one-electron functions $\psi_i(N)$, expressing the three-space component of the molecular orbital i of an electron with explicit coordinates x_i. This function is called the Hartree product $\Psi^{HF}(\mathbf{x}_1,...,\mathbf{x}_N)$.

Electronic wavefunctions are actually antisymmetric with respect to exchange (electrons are fermions), but the Hartree product does not satisfy the antisymmetry principle. These defects of the Hartree procedure were corrected by Slater (Slater, 1930), called the self-consistent-field-method (SCF). Slater uses determinant functions to provide the asymmetry of the wave function.

$$\Psi(\mathbf{x}_1,...,\mathbf{x}_N)=\frac{1}{\sqrt{N!}}\det\begin{pmatrix}\psi_1(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1)\\ \vdots & \ddots & \vdots\\ \psi_1(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N)\end{pmatrix} \tag{2.9}$$

which is called the Slater determinant, where $\mathbf{x}\equiv\{\mathbf{r},\sigma\}$ is the set of all spatial and spin coordinates of one electron. This form is not arbitrary, it is the simplest form that holds the antisymmetry principle for the wavefunction: when the coordinates of two electrons are exchanged in the wavefunction, the result equals the original wavefunction but with opposite sign. This principle is just a mathematical expression of Pauli's exclusion principle, which states that two electrons in a system cannot have identical quantum numbers.

The energy corresponding to the wavefunction can be calculated as:

$$E = \frac{\left\langle \Psi \left| \hat{H} \right| \Psi \right\rangle}{\left\langle \Psi \left| \Psi \right\rangle \right.} = \frac{\int \Psi^* \hat{H} \Psi \, d\tau}{\int \Psi^* \Psi \, d\tau} \tag{2.10}$$

where ψ^* is the complex conjugate of ψ, and the integration is with respect to three spatial coordinates and one spin coordinate, for each electron, symbolized by ($d\tau$).

The variational principle states that the energy calculated from Eq. (2.10) must be greater than or equal to the true ground state energy E_0.

$$E = \frac{\left\langle \Psi \left| \hat{H} \right| \Psi \right\rangle}{\left\langle \Psi \left| \Psi \right\rangle \right.} \geq E_0 = \frac{\int \Psi_0 \left| \hat{H} \right| \Psi_0}{\int \Psi_0 \left| \Psi_0 \right.} \tag{2.11}$$

Applying the theorem that the value of a determinant is unchanged by any non-singular linear transformation, we may choose the ψ to be an orthonormal set. We now introduce a Lagrange multiplier ε_i to impose the condition that the ψ are normalised, and minimise with respect to the ψ

$$\frac{\partial E}{\partial \psi_i} = 0 \tag{2.12}$$

They reduce to a set of one-electron equations of the form:

$$\hat{F} \psi_i = \varepsilon_i \psi_i \tag{2.13}$$

which is called the HF equation, where the Fock operator $\hat{F}$ is defined as

$$\hat{F} = -\frac{1}{2}\nabla^2 - \sum_{A=1}^{M} \frac{Z_A}{\left| r - R_A \right|} + \hat{v}^{HF} \tag{2.14}$$

The Fock operator, $\hat{F}$, describes the effect of all particles, including the other N-1 electrons, on the solution of one electron of the system. In the expression for the Fock operator the first term is the kinetic energy of the electron, the second is the interaction of the electron with the nuclei, and the third ($\hat{v}_i^{HF}$) is the exact-exchange operator representing the interaction of the electron with the rest of the

electrons in the system. This last operator is due to the effective reduction of the electrostatic repulsion between electrons of equal spin and the cancellation of a self Coulomb interaction of electrons, which means that no electron interacts with itself (Pisani C., 1996).

The Hartree Fock approximation describes the motion of one electron in the average field created by all other electrons of the system. The difference between the exact total energy and the total energy given by the HF method is generally called the correlation energy (despite the fact that some electronic 'correlation' is already included in the HF method). Although in some cases the contribution of the correlation energy is small and can be neglected, in many cases it plays an important role, and different post-HF methods have been developed to deal with the discrepancy. Most of these methods are based on relaxing the condition of the Slater form for the wavefunction, which is then expressed as a combination of different determinants including not only the occupied orbitals but also the empty ones. However, in the context of solid state calculations, a different approach is often employed to include correlation effects: the density functional theory method.

2.1.2 Density functional theory

The Density Functional Theory (DFT) represents an alternative approach to solving the many-electron quantum mechanical problem represented by equation 2.3. In DFT the central role is played by the so-called electron density $\rho(\mathbf{r})$, and not by the many-electron wavefunction $\Psi\left(\overline{x_1}, \overline{x_2}, ..., \overline{x_N}\right)$. In this section, after

having introduced the concept of electron density, we will state the two theorems that form the theoretical foundation of DFT, and the most popular computational implementation of this method. Finally, we will consider some common approximations of DFT in computational materials chemistry.

2.1.2.1 The Electron Density

The electronic density is defined as the following multiple integral over the spin coordinates of all electrons and over all but one of the spatial variables:

$$\rho(\bar{r}) = \mathrm{N}\int ...\int \left|\Psi\left(\overline{x_1}, \overline{x_2}, ..., \overline{x_N}\right)\right|^2 d\overline{x_1} d\overline{x_2} ... d\overline{x_N} \tag{2.15}$$

$\rho(\bar{r})$ determines the probability of finding any of the N electrons within the volume element $d\bar{r}_1$ but with arbitrary spin while the other N-1 electrons have arbitrary positions and spin in the state represented by Ψ.

$\rho(\bar{r})$ is a non-negative function of only three spatial variables which vanishes at infinity and integrates to the total number of electrons:

$$\rho\left(\vec{r} \to \infty\right) = 0 \tag{2.16}$$

$$\int \rho(\vec{r}) d\vec{r}_1 = \mathrm{N} \tag{2.17}$$

Unlike the wave function, the electron density is an observable and can be measured experimentally, e.g. by x-ray diffraction (Hajdu *et al.*, 1987).

2.1.2.2 The Hohenberg - Kohn Theorems

The density functional theory (DFT) is based on two theorems, proven by Hohenberg and Kohn in 1964 (Hohenberg & Kohn, 1964), and a computational scheme proposed by Kohn and Sham in the following year (Kohn & Sham, 1965).

The first theorem proved that the total external potential $V(\mathbf{r})$ acting on the electrons and the total number of electrons N are uniquely determined by the electronic density $\rho(\mathbf{r})$ of the ground state of the system. They also proved that $V(\mathbf{r})$ and N uniquely determine $\rho(\mathbf{r})$, as well as the Hamiltonian. Therefore, all properties of a physical system can be calculated if $\rho(\mathbf{r})$ is known and we can write the energy of the Schrödinger equation as a functional of the electron density:

$$E[\rho(r)] = T[\rho(r)] + V_{ee}[\rho(r)] + V_{ext}[\rho(r)] \tag{2.18}$$

$$F_{HK}[\rho(r)] = T[\rho(r)] + V_{ee}[\rho(r)] \tag{2.19}$$

where T is the total kinetic energy and V_{ee} is the functional describing the electron-electron interaction, together called the Hohenberg-Kohn fuctional

(F_{HK}), which is not dependent on the external potential, and V_{ext} is the Coulombic attraction between electrons and nuclei.

Hohenberg and Kohn also demonstrated a variational principle for the energy functional, which states that the density of the ground state ρ_0 minimizes the energy functional $E[\rho]$, and that the minimum of $E[\rho]$ is the energy of the ground state:

$$E[\rho_0] = E_0 \tag{2.20}$$

According to Hohenberg - Kohn theorems, the minimization of the functional:

$$E[\rho(r)] = F_{HK}[\rho(r)] + \int \rho(r) V_{ext}(r) dr \tag{2.21}$$

where $\rho(r)$ is a real density, which satisfies the necessary boundary conditions $\rho(r) > 0$ and $\int \rho(r) dr = N$ (equation 2.9 and 2.10), would give the ground state electron density $\rho_0(r)$, and knowledge of all the properties of the system. However, the exact form of the "universal" functional $F_{HK}[\rho(r)]$ (i.e. $T[\rho(r)]$ and $V_{ee}[\rho(r)]$) is not known.

2.1.2.3 The Kohn - Sham approach

The computational implementation of DFT is based on the approach developed by Kohn and Sham (Kohn & Sham, 1965). The method introduced a new fictitious system of N non-interacting electrons that has the same density as the real system

of interest, where the electrons do interact ($V_{ee}[\rho(r)]=0$). In this system the energy and the electron density are known from the one-electron wave function:

$$T_s[\rho]=\sum_{i=1}^{N}\left\langle \psi_i \left| -\frac{1}{2}\nabla^2 \right| \psi_i \right\rangle \tag{2.22}$$

The suffix (s) means that it is not the true kinetic energy, but that of a system of non-interacting electrons, where the density for the non-interacting electrons can simply be generated as:

$$\rho(\mathbf{r})=\sum_{i=1}^{N}\left|\psi_i\right|^2 \tag{2.23}$$

In a real system the electrons are interacting, and $T_s[\rho]$ does not provide the total kinetic energy. However, the difference between the exact kinetic energy and that calculated by assuming non-interacting orbitals is small.

The energy functional can be written as:

$$E[\rho(\mathbf{r})]=T_s[\rho(\mathbf{r})]+V_{ne}[\rho(\mathbf{r})]+J[\rho(\mathbf{r})]+E_{xc}[\rho(\mathbf{r})] \tag{2.24}$$

where $V_{ne}[\rho(\mathbf{r})]$ is the nuclear-electron interaction, $J[\rho(\mathbf{r})]$ is the classical Coulomb interaction between two charge distributions and the final term $E_{xc}[\rho(\mathbf{r})]$ is the exchange-correlation energy. It actually contains a correction to the kinetic energy deriving from the interacting nature of the electrons, and all

non-classical corrections to the electron – electron repulsion energy, including the exchange and correlation in the HF theory.

$$E_{xc}[\rho(\mathbf{r})] = \left(T[\rho(\mathbf{r})] - T_s[\rho(\mathbf{r})]\right) + \left(V_{ee}[\rho(\mathbf{r})] - J[\rho(r)]\right) \tag{2.25}$$

The form of the $E_{xc}[\rho(\mathbf{r})]$ functional is a critical aspect in the DFT. For such an approximate expression, one can apply the variational principle of the Kohn-Sham theory: it can be shown that a set of orbitals ψ_i that minimise the energy satisfy the equation:

$$\hat{h}^{KS}\psi_i = \varepsilon_i\psi_i, \tag{2.26}$$

which is analogous to the HF equation for one electron, now with the Kohn-Sham (KS) operator defined as:

$$\hat{h}^{KS} = -\frac{1}{2}\nabla^2 - \sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}-\mathbf{R}_A|} + \frac{1}{2}\int\frac{\rho(\mathbf{r}')}{\mathbf{r}-\mathbf{r}'}d\mathbf{r}' + V_{xc}(\mathbf{r}) \tag{2.27}$$

where

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} \tag{2.28}$$

is the functional derivative (analogous to a function derivative) of the exchange–correlation functional. The KS equation should be solved iteratively like the HF equation, since the KS operator itself depends on the orbitals χ_i of the solution.

The solution of the Kohn-Sham (KS) equation is exact, in contrast with HF theory; we only need to know the functional form of the exchange correlation energy term E_{xc}. Several different functionals have been proposed. And here, we shall briefly describe two important levels of approximation: the local density approximation (LDA) and the generalised gradient approximation (GGA).

2.1.2.4 Local Density Approximation (LDA)

The local density approximation (LDA) (Parr & Yang, 1989, Hohenberg & Kohn, 1964) refers to the uniform electron gas, where the local exchange-correlation energy per electron depends on the local charge density only:

$$E_{xc}^{LDA}[\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r})] = \int d\mathbf{r}\,\varepsilon_{xc}^{LDA}(\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r})) \tag{2.29}$$

where $\varepsilon_{xc}^{LDA}(\rho_\alpha,\rho_\beta)$ is the exchange – correlation energy per particle of a uniform electron gas (jellium) of the density ρ, which parameterisations have been proposed by Perdew-Zunger (Perdew *et al.*, 1981), and Vosko-Wilk-Nusair (Vosko *et al.*, 1980). The LDA is thus a first-principles approximation, in the sense that its formulation does not contain parameters which have been fitted empirically to calculated or experimental results. By construction, the LDA is exact for a uniform system, and a good approximation for slowly-varying systems like metals. Actually, the LDA still provides moderate accuracy for some systems where the density varies more rapidly over space, in particular in solid-state calculations, although it usually underperforms in the case of atoms and molecules.

In general, the exchange – correlation energy per particle ε_{xc} is a functional of the spin densities, that is, it depends on the values of the spin densities in the whole space and not only of their value at the integration point **r**, as in the LDA (which is the reason why this approximation is called local). Non-local effects in the functional can be included by incorporating a dependence of ε_{xc} on the gradients of the spin densities, which leads to the generalised gradient approximation (GGA).

2.1.2.5 Generalised Gradient Approximation Functionals (GGA)

The local character of the E_{xc} term in the LDA is corrected by the inclusion of terms that depend upon the gradient of the density, and not only on its value, at each point of the space:

$$E_{xc}^{GGA}[\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r})] = \int d^3\mathbf{r} f(\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r}),\nabla\rho_\alpha(\mathbf{r}),\nabla\rho_\beta(\mathbf{r})) \qquad (2.30)$$

Several approaches have been made to find expressions for the exchange-correlation functional, for example by Becke-Lee-Yang-Parr (B-LYP) (Becke, 1988, Lee *et al.*, 1988), Hampercht-Cohen-Tozer-Handy (HCTH) (Hamprecht *et al.*, 1998), Perdew-Wang (PWGGA) (Perdew & Yue, 1986, Perdew & Wang, 1992), Perdew 86 (P86) (Perdew, 1986), Perdew-Burke-Ernzerhof (PBE) (Perdew *et al.*, 1996). Expressions for the exchange-correlation functional are obtained via analytical functions, whose parameters are obtained either by fitting to experimental data (Becke, 1988, Lee *et al.*, 1988, Hamprecht *et al.*, 1998), or

derived from the principles of quantum mechanics (Perdew & Yue, 1986, Perdew & Wang, 1992, Perdew, 1986, Perdew *et al.*, 1996).

In our work we have used a GGA exchange correlation developed by Perdew, Burke and Ernzerhof (PBE). The authors presented the construction of a simplified GGA for exchange and correlation, in which all parameters are fundamental constants. The exchange correlation can be arbitrarily written as the sum of exchange and correlation functionals:

$$E_{xc}^{GGA}[\rho_\alpha(\mathbf{r}),\rho_\beta(\mathbf{r})]=\frac{1}{2}E_x^{GGA}[2\rho_\alpha(\mathbf{r})]+\frac{1}{2}E_x^{GGA}[2\rho_\beta(\mathbf{r})] \tag{2.31}$$

$$E_x^{GGA}(\rho)=\int d^3\mathbf{r}\rho\varepsilon_x^{unif}(\rho)]F_x(s) \tag{2.32}$$

where:

$$\rho=\rho_\alpha+\rho_\beta \tag{2.33}$$

$$\varepsilon_x^{unif}(\rho)=-3k_F/4\pi \tag{2.34}$$

$$k_F=\left(3\pi^2\rho\right)^{1/3} \tag{2.35}$$

$$s=\left|\nabla\rho\right|/2k_F\rho \tag{2.36}$$

The real-space cutoff gave a numerical function $F_x(s)$, which is:

$$F_x(s) = 1 + k - k\Big/\left(1 + \mu s^2/k\right) \tag{2.37}$$

where $\mu = 0.21951$ (to preserve the good LDA description of the exchange-correlation energy in the linear response of the uniform gas) and $k = 0.804$. Becke (Becke, 1988) proposed this form, but with empirical coefficients.

$$E_x^{GGA}(\rho_\alpha, \rho_\beta) = \int d^3 r \rho \varepsilon_x^{unif}(\rho) F_{xc}(r_s, \zeta, s) \tag{2.38}$$

Equation 2.25 represents any GGA exactly when $\zeta = \left(\rho_\alpha - \rho_\beta\right)\Big/\rho$ is independent of r, and is always approximately valid. The GGA proposed in this method retains correct features of the LDA, and combines them with the most energetically important features of gradient-corrected nonlocality.

In summary, the PBE functional yields great improvements on LDA for the total energies of atoms or atomization energies of molecules (Perdew *et al.*, 1993) but has a mixed history of successes and failures for solids (Garcia *et al.*, 1992, Singh & Ashkenazi, 1992, Ozolins & Korling, 1993, Filippi *et al.*, 1994, Hamann, 1996).

2.1.3 Basis sets

One of the approximations in all quantum mechanical calculations is the introduction of the basis set, i.e. the molecular orbitals which are formed as a linear combination of atomic orbitals (LCAO):

$$\psi_i = \sum_{\nu=1}^{K} c_{\nu i}\phi_\nu \tag{2.39}$$

The one-electron wave orbitals ϕ_ν are called basis functions and often correspond to the atomic orbitals, where the index ν categorizes different functions used, and therefore only the expansion coefficients $c_{\nu i}$ are varied. For a given basis set and a given functional form of the wavefunction the best set of the coefficients is that for which the energy is a minimum, at which point

$$\frac{\partial E}{\partial c_{\nu i}} = 0 \tag{2.40}$$

for all coefficients $c_{\nu i}$. The objective is thus to determine the set of coefficients that gives the lowest energy for the system. In order to describe accurately the solution, an infinite number of basis functions can be used, but including more basis functions also increases the computational cost. Hence, the right balance between the accuracy and cost depends on the right choice of the representative basis set.

Calculations of interaction energies are susceptible to basis set superposition error (BSSE) if they use finite basis sets. As the atoms of interacting molecules (or of different parts of the same molecule) or two molecules approach one another, their basis functions overlap.

The program SIESTA (Soler *et al.*, 2002) for example, uses numerical bases for the calculations. Numerical basis functions are values on atomic-centered spherical-polar mesh (Altmann *et al.*, 1996), rather than as analytical functions (i.e., Gaussian orbitals). The angular portion of each function is the appropriate spherical harmonic $Y_{lm}(\theta,\varphi)$. The radial portion $F(\mathbf{r})$ is obtained by solving the atomic DFT equations numerically. A reasonable level of accuracy is usually obtained by using about 300 radial points from the nucleus to an outer distance of 10 bohr (~5.3 Å). The particular type of numerical basis selected was suggested our calculations to be the most efficient for fast-convergent 3D numerical integrations.

In order to reduce the computational cost, basis sets use the so-called pseudopotential approximation to describe core states.

2.1.4 Pseudopotential

The fundamental idea of the pseudopotential is the replacement of the strong Coulomb potential of the nucleus and the effects of the tightly bound core electrons by an effective ionic potential acting on the valence electrons. The pseudopotential is a function that gives wavefunctions with the same shape as the true wavefunction outside the core region but fewer nodes inside the core region, as illustrated in Figure 2.1. It can be generated in an atomic calculation and then used to compute properties of valence electrons in molecules or solids, since the core states remain almost unchanged. This procedure works well because most

physical properties of solids are dependent on the valence electrons to a much greater extent than on the core electrons.

In the region occupied by the core electrons ($r < r_c$ in the Figure) the valence wavefunctions oscillate rapidly due to the strong Coulomb potential in this region. These oscillations maintain the orthogonality between the core wavefunctions and the valence wavefunctions, which is required by the exclusion principle. In the so-called norm-conserving pseudopotential (NCPP) approach (Hamann *et al.*, 1979), the all-electron wavefunction is replaced by a soft nodeless pseudo-wavefunction inside the core radius r_c, with the crucial restriction that the pseudo wavefunction must have the same norm (electronic charge) as the all-electron wavefunction within the chosen core radius. For $r > r_c$ the pseudo and the all-electron wavefunction are identical.

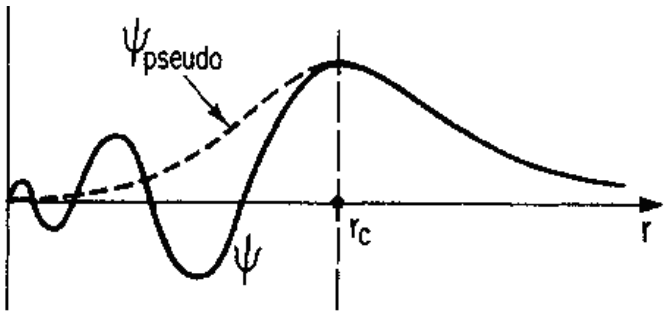


Figure 2.1 Schematic representation of pseudopotential.

The use of these pseudopotentials has the effect of reducing the number of terms required for the plane wave expansion of the wavefunction, which in turn drastically reduces the scale of the computational problem.

2.1.5 Geometry Optimisation

When modelling chemical systems, be they molecules or solid, we are especially interested in minimum points on the energy surface. Minimum energy arrangements of the atoms correspond to a stable state of the system; any movement away from a minimum gives a configuration with a higher energy. At a minimum point the first derivative of the energy function with respect to all nuclear coordinates is zero and the second derivates are all positive:

$$\frac{\partial E}{\partial r_i} = 0\,;\ \frac{\partial^2 E}{\partial^2 r_i} > 0 \tag{2.41}$$

There are many different algorithms that can be used to find a minimum in a function of several variables: those which use derivatives of the energy with respect to the coordinates and those which do not (Gill *et al.*, 1981). In our SIESTA calculations we have used the conjugate gradients method, which requires not only the evaluation of the function E, but also of its gradient ∇E. Each step is taken in the direction for which E decreases the fastest, which means that the first step is in the direction of $-\nabla E$. The conjugate gradients method moves in a direction d_k from point r_k where d_k is computed from the gradient at the point of the previous direction vector d_{k-1}:

$$d_k = \Delta E_k + \gamma_k d_{k-1} \tag{2.42}$$

γ_k is a scalar constant given by

$$\gamma_k = \frac{\Delta E_k \cdot \Delta E_k}{\Delta E_{k-1} \cdot \Delta E_{k-1}} \tag{2.43}$$

The idea of this method is to choose each new step in a direction that is conjugate to the directions used in the previous steps, so as to avoid undoing any of the minimization work done in previous steps.

2.2 Atomistic Simulation

Classical simulations are based on a number of assumptions, which together form what is usually called the Born model of solids (Born & Huang, 1954). The Born model of solids assumes that the ions in the crystal interact via long-range electrostatic forces and short-range forces, including both repulsions and van *der* Waals attractions between neighbouring electron charge clouds, which can be described by the equation:

$$E(r_{ij}) = \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon_0 (r_{ij} + 1)} + \sum_{ij} \Phi_{ij}(r_{ij}) \tag{2.44}$$

The first term equation 2.44 is the long-range electrostatic interaction between the charges of all the ions, where q_i and q_j are the charges on atoms i and j respectively, r_{ij} is the distance between atoms i and j, and ε_0 is the permitivity of free space. The second term is the short-range van *der* Waals forces between neighbouring electron clouds, with the potential functions used to describe these forces below. Unfortunately, the electrostatic part of the potential (the first term of equation 2.44) converges very slowly in real space. This problem is overcome by

use of the Ewald method (Ewald, 1921), which divides the electrostatic sum into two parts: the first (at short distance) is summed in real space, while the second (at longer distance) is computed in reciprocal space.

2.2.1 The Ewald Method

The Ewald summation was introduced in 1921 to calculate the long-range interactions between particles and their infinite periodic images. This method is based on splitting the conditionally convergent series into two convergent series, one in the real space and one in the reciprocal space, each of which converges rapidly. Each charge is considered to be surrounded by a neutralising charge distribution of equal magnitude but opposite sign, as shown in the Figure 2.2 (a). A Gaussian charge distribution of the following functional forms is commonly used:

$$\rho_A(\mathbf{r}) = \frac{q_A \alpha^3}{\pi^{3/2}} \exp(-\alpha^2 \mid \mathbf{r} - \mathbf{R}_A \mid^2) \tag{2.45}$$

It can be proved that the electrostatic energy of the sum of the point charges plus the screening Gaussian densities is now:

$$U_{real}^{Coulomb} = \frac{1}{2} \sum_{A,B} q_A q_B \sum_{\mathbf{T}}{}' \frac{\mathrm{erfc}(\alpha \mid \mathbf{R}_{AB} + \mathbf{T} \mid)}{\mid \mathbf{R}_{AB} + \mathbf{T} \mid} \tag{2.46}$$

where $erfc(x) = \frac{2}{\pi^{1/2}} \int_x^\infty \exp(-t^2) dt$ is the complementary error function. This is a rapidly convergent series which can be evaluated and truncated in real space.

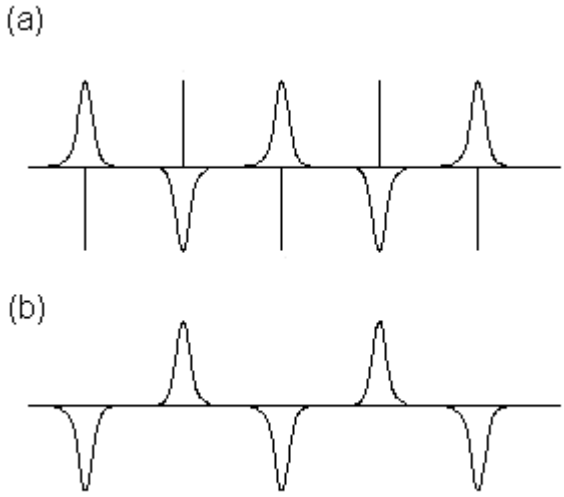


Figure 2.3 Charge distribution in the Ewald sum. The energy of distributions a) is calculated in real space and b) in reciprocal space.

Another Gaussian charge distribution must be added in order to maintain the neutrality of charges (see Fig. 2.3b). This second distribution gives an electrostatic energy that is easier to calculate in reciprocal space:

$$U_{reciprocal}^{Coulomb} = \frac{1}{2} \sum_{\mathbf{G} \neq \mathbf{0}} \sum_{A,B} \frac{q_A q_B}{\pi V_{cell}} \frac{4\pi^2}{G^2} \exp(-G^2 / 4\alpha^2) \cos(\mathbf{G} \cdot \mathbf{R}_{AB}) \qquad (2.47)$$

where **G** are the reciprocal vectors and V_{cell} is the volume of the unit cell of the crystal. We should finally subtract the spurious interaction of each Gaussian density with itself, which is included in the real space sum:

$$U_{self-energy}^{Coulomb} = \frac{\alpha}{\pi^{1/2}} \sum_{A} q_A^2 \tag{2.48}$$

The total electrostatic energy per unit cell can then be calculated as:

$$U^{Coulomb} = U_{real}^{Coulomb} + U_{reciprocal}^{Coulomb} - U_{self-energy}^{Coulomb} \tag{2.49}$$

2.2.2 Interatomic Potential functions

The total short-range energy is usually expressed as:

$$U^{short-range}(r_1,...,r_N) = \sum_{A,B} U_{AB}(r_A, r_B) + \sum_{A,B,C} U_{ABC}(r_A, r_B, r_C) + ... \tag{2.50}$$

where U_{AB} refers to the two-body interactions, U_{ABC} to three-body interactions, and so on. This expression combines a number of components, including non-bonded interactions (repulsion and van der Waals attraction), electronic polarisability and, covalent interactions, modelled by bond-bending and bond-stretching terms.

For ionic or semi-ionic solids, the most frequently used functional form for the short-range two-body potential is the so-called Buckingham potential:

$$U_{AB}^{buck}(r_{AB}) = A\exp\left(-r_{AB}/\rho\right) - \frac{C}{r_{AB}^6} \tag{2.51}$$

where the first term describes the Pauli repulsive interaction and the second term represents the van der Waals interactions. As a result of this opposite interactions, the Buckingham potential exhibits an energy minimum (Figure 2.4).

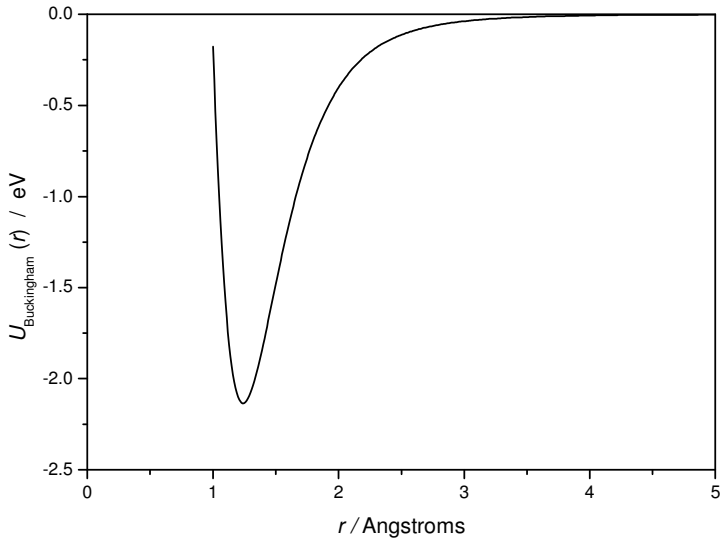


Figure 2.4 short-range interatomic interactions in the form of a Buckingham potential

For these systems, the Buckingham potential generally performs better than the simpler Lennard-Jones potential:

$$U_{AB}^{LJ}(r_{AB}) = \frac{B}{r_{AB}^{n}} - \frac{C}{r_{AB}^{m}} \tag{2.52}$$

where B and C are adjustable parameters and n and m are integer numbers, generally n=12, m=6. In both potentials, the first term models the repulsive interactions due to the Pauli forces, and the second one represents the attractive interactions due to the van der Waals dispersion forces. At short distances the repulsive forces become much higher than the attractive ones.

For a description of interactions between atoms bonded by a partially covalent bond, a Morse potential can be used

$$U_{ij}^{morse} = A_{ij}\left[\left\{1-\exp\left[k_{ij}\left(r_{ij}-r_0\right)\right]\right\}^2 - 1\right] \tag{2.53}$$

where A_{ij} is the dissociation energy of the bond between the two atoms *i* and *j*, r_{ij} is the bond length, r_0 is the equilibrium bond length, and k_{ij} is a function of the slope of the potential energy well which can be determined from spectroscopic data.

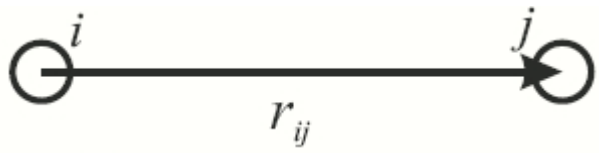


Figure 2.5 Interatomic bond vectors

A simpler alternative to the Morse potential is the harmonic potential:

$$U_{ij}^{harm} = \frac{1}{2}k_{ij}\left(r_{ij}-r_0\right)^2 \tag{2.54}$$

where k_{ij} is the harmonic force constant.

The second main contribution to the short-range energy is the three-body energy, which is often necessary for the description of covalent and semi-covalent bonds,

because of the directionality of these interactions. It is usually expressed as a harmonic potential:

$$U_{ijk}^{harmonic}(\theta_{ijk}) = \frac{1}{2} k_{ijk}(\theta_{ijk} - \theta_0)^2 \tag{2.55}$$

where k_{ijk} is a bond-bending force between bonds *ij* and *ik,* and θ_o is the equilibrium angle between the three atoms.

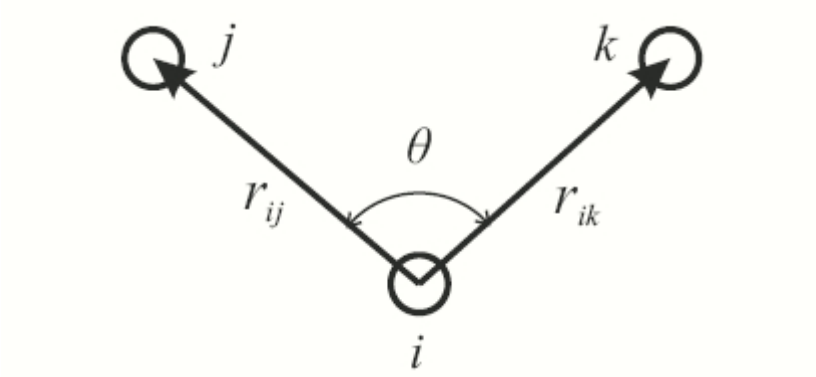


Figure 2.6 The valence angle and associated vectors

Four-body interactions, as well as any other interaction involving more than four centres, make a very small contribution to the total energy of ionic or semi-ionic systems, so usually they are only used in particular molecules, *e.g.* amino acids, and proteins, where the geometry of the molecule is flexible and complex. Two planes are defined, between the atoms, *i*, *j* and *k* and between *j*, *k* and *n* with the potential energy being dependent on n times the angle between these planes, shown in Figure 2.7, thus giving rise to a minimum when the angle is zero, and multiples of $2\pi/n$ radians.

$$U_{ijkn}^{four-body}(\Phi_{ijkn}) = K_{ijkn}\left[1-\cos\left(n\Phi-\gamma\right)\right] \tag{2.56}$$

where K_{ijkn} is the torsional force constant, γ is phase offset in degrees, and n is the periodicity of the torsion.

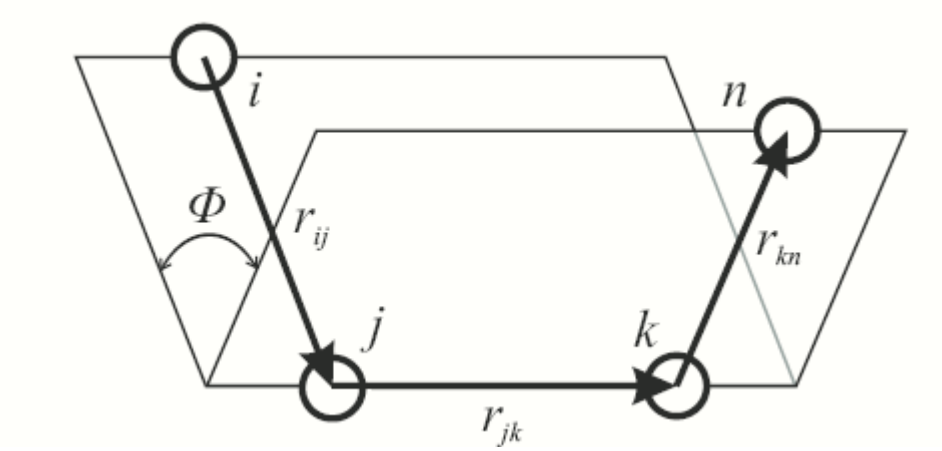


Figure 2.7 The dihedral angle and associated vectors

Short-range interactions tend to zero rapidly when the interatomic distance increases. Therefore we can assume that every atom interacts not with all the atoms, but only with those atoms inside a certain spherical region of radius r_c. This cut-off radius must be large enough to ensure that the interactions with the atoms outside this region are indeed negligible.

2.2.2.1 Electronic Polarisability

In the model we have described so far every atom is modelled by a point charge, without any dipole moment. Although this may be an acceptable approximation in some cases, there are many situations where polarisability plays a crucial role, for example when defects or surfaces are present.

The electronic polarisability of the atoms in the hydroxyapatite material is included via the shell model of Dick and Overhauser (Dick & Overhauser, 1958), where each polarisable ion, here oxygen, is represented by a core and a massless shell, connected by a spring. This simple *shell model* for the atomic polarisability is illustrated in Figure 2.8.

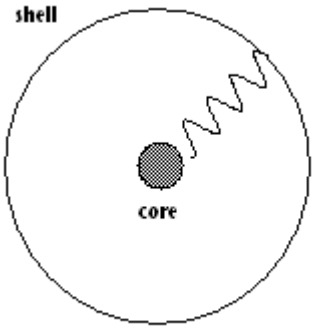


Figure 2.8 Schematic representation of an ion composed by a core and a shell.

The spring constant and the charges of the core and shell then determine the polarisability of the atom:

$$\alpha = \frac{Y^2}{k} \tag{2.57}$$

where k is the force constant for the spring connecting the core and the shell in the spring, and Y is the charge on the shell. Both are fitted by empirical fitting to experimental properties such as dielectric and elastic constants. The sum of the core charge and the shell charge is the total charge of the ion.

The electronic polarisability is slightly different for molecular dynamics simulations, such as DL_POLY (Smith & Forester, 1996). The general principle is the same but for molecular dynamics simulations the shell is given a mass, which is negligibly small compared to the mass of the core, to ensure that there is no exchange of energy between the core and shell during the simulation.

2.2.3 Molecular Dynamics

Molecular dynamics simulations (MD) are a method to simulate time-dependent behaviour of a polyatomic system. All forms of molecular dynamics simulations are based on the behaviour of the molecular systems according to Newton's laws of motion, in particular his second law, where the force applied to a system is proportional to its acceleration, or variation in speed over time.

$$m_i \frac{\partial^2 \vec{r_i}}{\partial t^2} = F_i\left(\vec{r_1}, \vec{r_2}, \ldots \vec{r_N}\right) \tag{2.58}$$

This equation describes the motion of a particle of mass m_i along the vector position $\vec{r_i}$ with $F_i\left(\vec{r_1}, \vec{r_2}, \ldots \vec{r_N}\right)$ being the force on the particle in that direction, from a potential *V* created by their interaction with all other particles of the system in an instant of time Δt.

$$F_i = \nabla_i V\left(r_1, r_{2,} \ldots r_N\right) \tag{2.59}$$

Under the influence of a continuous potential the motions of all the particles are coupled together, giving rise to a many-body problem that cannot be solved analytically. The equations of motion are then solved in a finite difference approximation at successive time steps (Δt), with the velocities and positions of the particle calculated each time step used for successive interactions.

2.2.3.1 Finite Difference Methods

There are many algorithms for integrating the equation of motion using finite difference methods, several of which are commonly used in molecular dynamics calculations. All algorithms assume that the positions and dynamical properties (velocities, accelerations, etc.) can be approximated as Taylor expansion series:

$$r(t+\Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) + \frac{1}{6}\Delta t^3 b(t) + \frac{1}{24}\Delta t^4 c(t) + \ldots \quad (2.60)$$

$$v(t+\Delta t) = v(t) + \Delta t a(t) + \frac{1}{2}\Delta t^2 b(t) + \frac{1}{6}\Delta t^3 c(t) + \ldots \quad (2.61)$$

$$a(t+\Delta t) = a(t) + \Delta t b(t) + \frac{1}{2}\Delta t^2 c(t) + \ldots \quad (2.62)$$

$$b(t+\Delta t) = b(t) + \Delta t c(t) + \ldots \quad (2.63)$$

where r and v are the positions and the velocities respectively, a is the acceleration, and b is the third derivate of the position with respect to time. The Verlet algorithm (Verlet, 1967) is probably the most widely used method for integrating the equations of motion, which uses the positions and accelerations at time *t*, and the position from the previous step, $r(t-\Delta t)$, to predict the new

positions at $t+\Delta t$, $r(t+\Delta t)$. The idea for this integration algorithm is based on a 3rd-order Taylor expansion series of the position:

$$r(t+\Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 + \frac{1}{6}b(t)\Delta t^3 + O(\Delta t^4) \tag{2.64}$$

$$r(t-\Delta t) = r(t) - v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 - \frac{1}{6}b(t)\Delta t^3 + O(\Delta t^4) \tag{2.65}$$

where $O(\Delta t^4)$ is the error associated with the truncation of the expression at the 3rd-order.

Then, the addition of 2.64 and 2.65 gives

$$r(t+\Delta t) = 2r(t) - r(t-\Delta t) + \Delta t^2 a(t) \tag{2.66}$$

and the velocities can be calculated from:

$$v(t) = \left[r(t+\Delta t) - r(t-\Delta t) \right] / 2\Delta t \tag{2.67}$$

However, this equation is subject to errors in Δt^2, and this may lead to a loss of precision in the calculation of the velocities compared with the position. Several variations on the Verlet algorithm have been developed. One such modification is called the leap-frog algorithm (Hockney, 1970), which uses the algorithm:

$$r\left(t+\Delta t\right)=r\left(t\right)+\Delta tv\left(t+\frac{1}{2}\Delta t\right) \tag{2.68}$$

$$v\left(t+\frac{1}{2}\Delta t\right)=v\left(t-\frac{1}{2}\Delta t\right)+\Delta ta\left(t\right) \tag{2.69}$$

where the velocities $v\left(t+\frac{1}{2}\Delta t\right)$ are first calculated from the velocities at time $\left(t-\frac{1}{2}\Delta t\right)$ and the accelerations at time *t*. The positions $r\left(t+\Delta t\right)$ are then deduced from the velocities just calculated together with the positions at time $r\left(t\right)$ using Eq. 2.68. The velocities at time *t* can be calculated as the average of the velocities at $\left(t+\frac{1}{2}\Delta t\right)$ and $\left(t-\frac{1}{2}\Delta t\right)$.

Two advantages of the leap-frog algorithm over the Verlet algorithm are that the leap-frog algorithm also includes the velocity and it does not need the calculation of the differences of large numbers. However, it is not possible to calculate the kinetic energy contribution to the total energy at the same time as the positions are defined, because the positions and velocities are not synchronised.

An alternative Verlet algorithm that does store positions, velocities and accelerations at the same time, which minimises rounding errors, has also been proposed (Swope *et al.*, 1982). This velocity-Verlet algorithm has the form:

$$r\left(t+\Delta t\right)=r\left(t\right)+\Delta tv\left(t\right)+\frac{1}{2}\Delta t^{2}a\left(t\right) \tag{2.70}$$

$$v(t+\Delta t) = v(t) + \frac{1}{2}\Delta t\left[a(t) + a(t+\Delta t)\right] \tag{2.71}$$

This algorithm has three stages. In the first stage it requires values of position (r), velocity (v) and acceleration (a) at time t and $(t+\Delta t)$. The first stage is to advance the velocities to $\left(t+\frac{1}{2}\Delta t\right)$ by integration of the force and then to advance the positions to a full step $(t+\Delta t)$ using the new half-step velocities.

Alternative algorithms are the Beeman algorithm and the Predictor-Corrector method. The reader is directed to the references (Beeman, 1976, Gear, 1971) for more details about these methods, which are more complex and time-consuming than the Verlet algorithm.

The choice of the time step, (Δt), in molecular dynamics simulation is also important. If the time intervals, (Δt), are too large (multiple fs), the MD simulations can explore more of the system configurations. However, the integrated equations for the displacement require that the force remains constant at each step, and the molecular vibrations need to be sampled at small enough intervals to ensure that they are sampled properly.

The magnitude of the time interval Δt is therefore limited by the maximum frequency, $\gamma_{\max}$ of the vibrations between the component particles, and must be

much smaller than $\gamma_{\max}$ in order to sample the fastest vibrational motion during the simulation:

$$\Delta t \ll \frac{1}{\gamma_{\max}} \tag{2.72}$$

At short Δt, the higher order predictor-corrector methods present minor fluctuations in the total energy of the Verlet. With intermediate Δt the fluctuations in both methods are similar and with big Δt the Verlet algorithm is a better quality.

2.2.3.2 Periodic Boundary Conditions

Periodic boundary conditions are applied to finite particle systems so that they can mimic an infinite system. The simulation box is replicated throughout space to form an infinite lattice by rigid translation in all three cartesian directions, completely filling the space.

In Figure 2.9, the shaded box represents the system we are simulating, while the surrounding boxes are exact copies in every detail - every particle in the simulation box has an exact duplicate in each of the surrounding cells. Even the velocities (indicated by the arrows) are the same. This arrangement is repeated infinitely to fill the whole of space. A result, whenever an atom leaves the simulation cell, it is replaced by another with exactly the same velocity, entering from the opposite cell face. So the number of atoms in the cell is conserved. Furthermore, no atom feels any surface forces, as these are now completely removed.

In the Figure r_{cut} is the cutoff radius that is normally applied when calculating the force between two atoms. As you can see, an atom may interact with one in the neighbouring cell (which is an image of one of the atoms in the simulation cell) because it is within the cutoff radius. It ignores the equivalent atom in the simulation cell because it is too far away. In other cases the interaction comes from an atom in the simulation cell itself. Thus the interaction that is calculated is always with the closest image. This is known as the *minimum image* convention. Note that the cutoff radius is always chosen so that an atom can interact with *only one image* of any given atom. This means that r_{cut} cannot be greater than half the width of the cell, i.e. $L_{box}>2r_{cut}$.

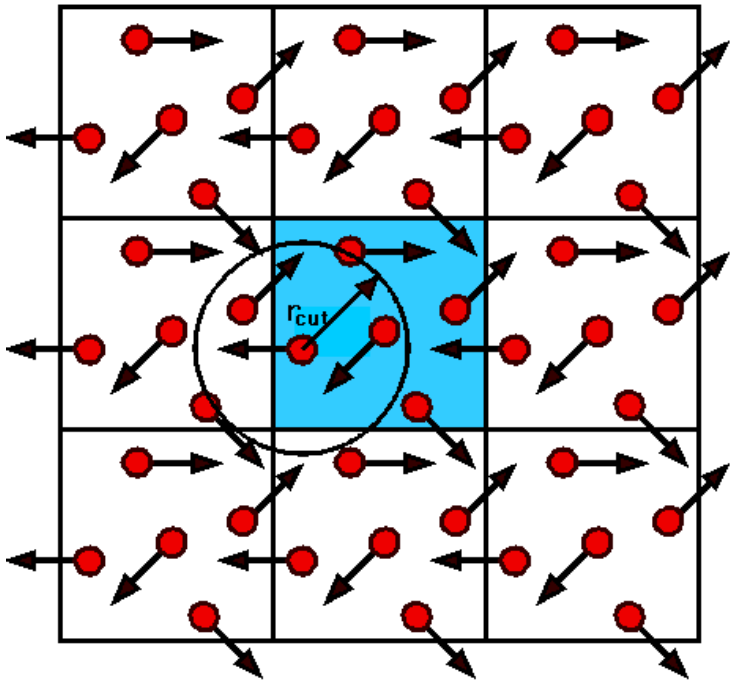


Figure 2.9 Schematic representation of periodic boundary conditions.

In the case of surface simulations, ideal model would be a slice of material that is infinite in two dimensions, but finite along the surface normal. In order to

accomplish this, it may seem natural to take advantage of periodic boundary conditions in two dimensions, but not the third. There are codes in which this technique is implemented, but it is more common to study a surface using a code that applies periodic boundary conditions in all three dimensions, and it is this approach we will discuss.

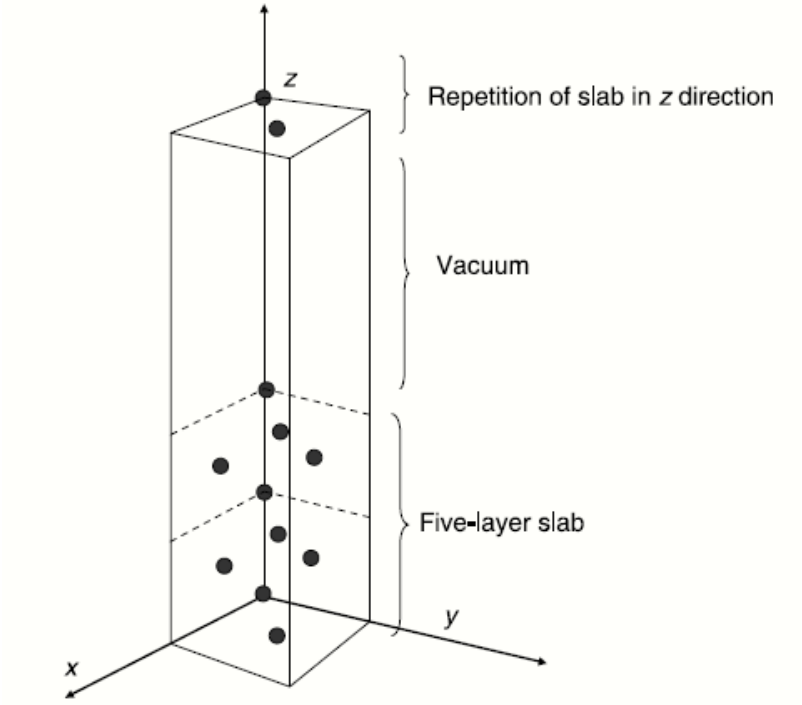


Figure 2.10 Schematic representation of supercell that defines a material with a solid surface when used with periodic boundary conditions in all three directions.

The basic idea is illustrated in Figure 2.10, where the supercell contains atoms along only a fraction of the vertical direction. The atoms in the lower portion of the supercell fill the entire supercell in the *x* and *y* directions, but empty space has been left above the atoms in the top portion of the supercell. This model is called a slab model since, when the supercell is repeated in all three dimensions, it defines a series of stacked slabs of solid material separated by empty spaces. The

empty space separating periodic images of the slab along the z direction is called the vacuum space.

2.2.3.3 Steps of the Dynamics Simulations

The first step in molecular dynamics simulations is to allow the system to come to equilibrium at a given temperature and pressure before data are collected. Each particle is located in accordance with a geometry test and its velocities are periodically scaled to return the system to the temperature input. Initially the velocity of each of the particles is standardized to achieve a linear time or resulting net zero, thus avoiding unnecessary displacement of the whole system. The initialization step is terminated when the temperature stabilizes (the energy of the system is constant) and no longer needs renormalization for this purpose because it has reached steady state.

The next step is to carry out the simulation in a given virtual time, which contains a significant number of steps Δt . In this process the positions and velocities of all particles in selected steps are recorded and stored. The conditions of the simulation, or ensemble, can have a number of forms:

- The microcanonical ensemble (NVE): constant number of particles, constant energy and constant volume,
- The canonical ensemble (NVT): a constant number of particles, constant volume and constant temperature,

- The isothermal-isobaric ensemble (NPT): a constant number of particles, constant pressure and constant temperature,
- The grand-canonical ensemble (μVT): a constant chemical potential, constant volume and constant temperature.

Molecular dynamics is traditionally performed under conditions of constant number of particles, temperature and volume, but its trajectory can be modified to sample from other ensembles.

Even in a constant *NVE* simulation it is usually desired to perform a simulation at a particular temperature. In fact, it is common practice during the equilibration phase to adjust the temperature to the desired value.

The instantaneous value of the temperature is related to the kinetic energy via the particles' velocities:

$$K(t) = \sum_{i=1}^{N} \frac{1}{2} m_i \left| \vec{v}_i(t) \right|^2 = \frac{3}{2} N k_B T(t) \tag{2.73}$$

where k_B is the Boltzmann's constant. Therefore, an obvious way to alter the temperature of the system is by rescaling velocities. In the *velocity Verlet* algorithm this may be achieved by replacing the equation

$$\vec{v}_i\left(t + \Delta t/2\right) = \vec{v}_i(t) + \frac{1}{2}\vec{a}_i(t)\Delta t \tag{2.74}$$

with

$$\vec{v}_i\left(t+\Delta t/2\right)=\sqrt{\frac{T_0}{T(t)}}\vec{v}_i(t)+\frac{1}{2}\vec{a}_i(t)\Delta t \tag{2.75}$$

where T_0 is the desired temperature, and $T(t)$ is the "instantaneous temperature" obtained from equation 2.74. Such modification means that we are no longer following Newton's equations, and the total energy is no longer conserved. Important data should not be collected in this stage: these "controlled temperature" simulations should be used only to bring the system from one state to another.

When the system has achieved equilibrium then the production phase can commence. At the start of the production phase all counters are set to zero and the system is allowed to evolve. In a microcanonical ensemble (*NVE*) no velocity scaling is performed during the simulation and so the temperature becomes a calculated property of the system. Various properties are routinely calculated and stored during the production phase for subsequent analysis and processing. Careful monitoring of these properties during the simulations can show whether the system is 'well behaved' or not; it may be necessary to restart a simulation if problems are encountered. It is also usual to store the positions, forces and velocities at regular intervals, from which other properties depending on these quantities can be determined at the end of the simulation.

2.2.3.4 Properties from Molecular Dynamics Simulations

A wide variety of properties can be calculated from molecular dynamics simulations, including the radial distributions functions, and diffusion coefficients.

The radial distribution functions provide valuable information of the structure of the system. The radial distribution function, $g(r)$, is the probability of finding an atom at a distance *r* from another atom.

$$g(r) = \frac{V}{N}\left\langle \frac{n(r)}{4\pi r^2 \Delta r} \right\rangle \tag{2.76}$$

where *V* is volume, $n(r)$ is the average number of atoms separated at a distance between *r* and $r + \Delta r$ from a given particle, and *N* is number of particles per unit volume.

Another property that is best studied by molecular dynamics is diffusion. Diffusion coefficients are obtained from the mean square displacement, MSD, defined as

$$MSD(t) = \left\langle \left| \vec{r}(t) - \vec{r}(0) \right|^2 \right\rangle \tag{2.77}$$

where $r(0)$ is the initial position of the molecule and $\vec{r}(t)$ is the position at time *t* via the Einstein relationship:

$$D = \lim_{x \to \infty} = \frac{\left\langle \left| \vec{r}(t) - \vec{r}(0) \right|^2 \right\rangle}{6t} \tag{2.78}$$

This is a good measure of the translational mobility of particles.

Chapter 3 Methodology

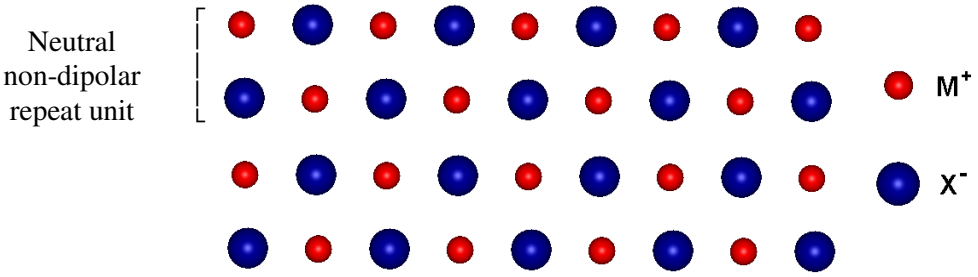
The goal of this thesis is to investigate computationally the interactions of the predominant molecules in the collagen protein with the apatite mineral, using a combination of computational techniques, density functional theory (DFT) and molecular dynamics simulations based on interatomic potentials. In this chapter we shall describe the methodologies applied in our work.

3.1. Surface Simulations

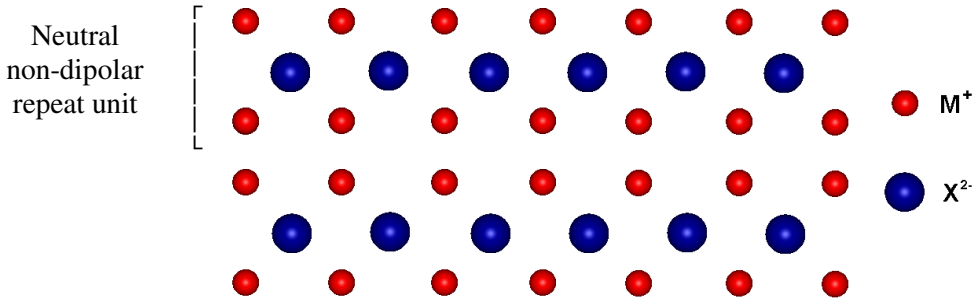
We have created the surface models using the METADISE program (Watson *et al.*, 1996), which is designed to model dislocations, interfaces and surfaces. We have followed the approach of Tasker (Tasker, 1979) where the simulation model consists of a series of charged planes parallel to the surface and periodic in two dimensions. Tasker classified surfaces in three different types:

1. Type I surfaces, where the repeat unit is a charge neutral stoichiometric layer (Figure 3.1(a)).
2. Type II surfaces, which comprise charged layers but in such a way that there is no dipole moment perpendicular to the surface (Figure 3.1(b)).
3. Type III surfaces where there is a dipole moment perpendicular to the surface (Figure 3.1(c)).

a) Type I



b) Type II



c) Type III

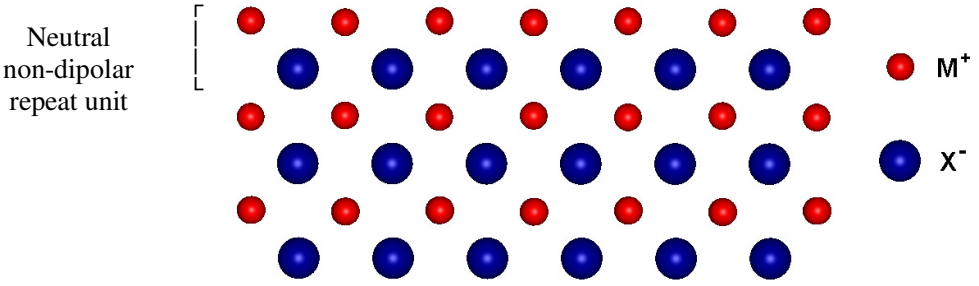


Figure 3.1 The three different types of surfaces as classified by Tasker

Type III surfaces are made up of stack of alternately charged planes, producing a dipole moment perpendicular to the surface. In this case, the surface model needs to be reconstructed to remove the dipole. One way of achieving this is to remove half the ions from the surface layer at the top of the repeat unit and transfer them to the bottom (Figure 3.2).

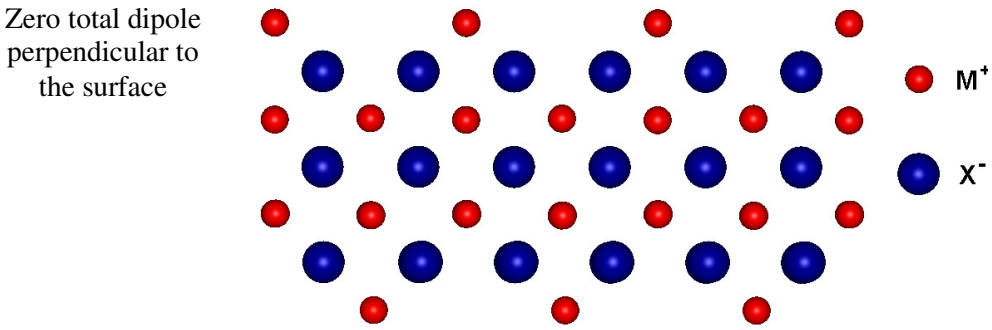


Figure 3.2 Stacking sequences showing a reconstructed Type III surface where half the surface ions have been shifted to the bottom of the unit cell, which has removed the dipole in the reaped unit.

3.1.1 Classification of surfaces by Miller Indices

There are many different planes along which it is possible to cleave the crystal and these different directions reveal different kinds of arrangement of the atoms on the surface. The orientation of the crystal plane is usually defined by three indices enclosed in parentheses (*hkl*), which are called Miller indices. The Miller indices are defined by the reciprocals of the intercepts on the axes in terms of the lattice constants a_1, a_2, a_3, which are then reduced to three integers (Figure 3.3). If a plane cuts an axis on the negative side the origin, the corresponding index is negative, denoted with an overbar, e.g. $(h\bar{k}l)$. The indices (*hkl*) may denote one plane or a set of parallel planes. For a cubic crystal it is not necessary to distinguish between (100), (010), (001), $(\bar{1}00)$, $(0\bar{1}0)$, and $(00\bar{1})$ surfaces, which all have the same structure but this is not generally true in other crystal systems.

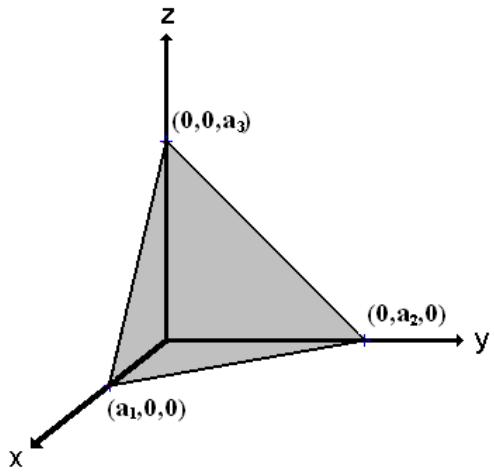


Figure 3.3 This plane intercepts the a_1, a_2, a_3.

We now discuss the crystal structure of interest for our study of hydroxyapatite, which is the hexagonal crystal system. For hexagonal crystals, a three-axis system to describe the planes is unsatisfactory, as shown in Figure 3.4. The two planes highlighted in Figure 3.4 are equivalent by symmetry, and yet their Miller indices do not show this relationship. For this reason, Miller indices are determined using a four-axis system, shown in Figure 3.5, and therefore four indices written as (*hkil*).

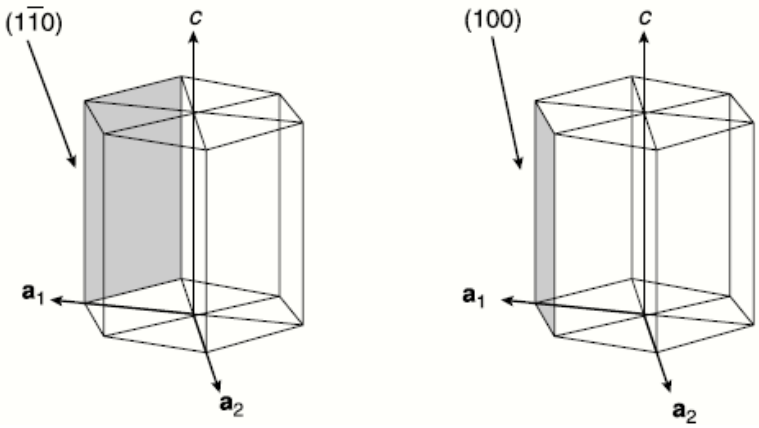


Figure 3.4 Labelling of planes in a solid using a three axis system. These two planes are equivalent by symmetry, but the indices obtained using three axes are not.

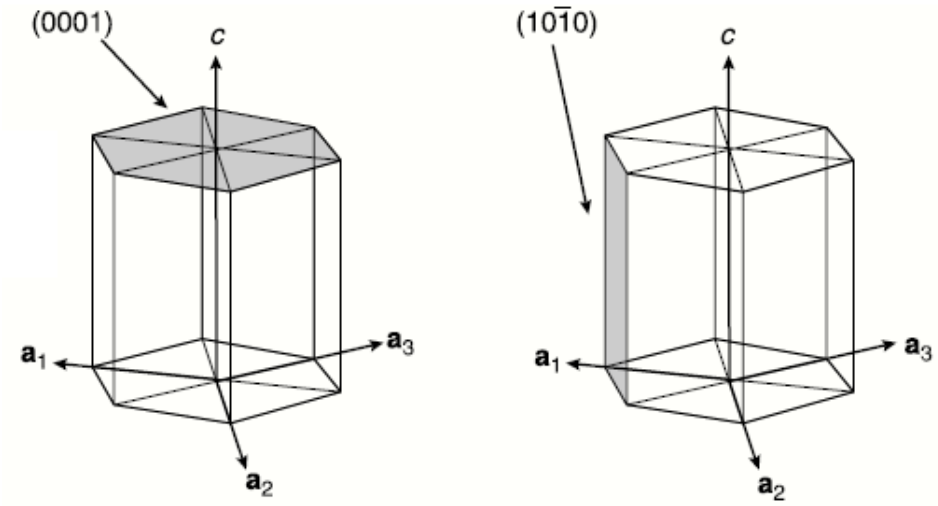


Figure 3.5 Miller indices for hexagonal crystal systems using the four-axis index.

3.1.2 Surface and Adsorption Energies

The surface energy is an inverse measure of the thermodynamic stability of the surface, and experimentally is defined as the energy required to cleave the bulk crystal exposing the surface. We can calculate surface energies using the expression:

$$\gamma = \frac{\left(E_{surf} - E_{bulk}\right)}{A} \tag{3.1}$$

where E_{surf} is the energy of the simulation cell containing the surface, E_{bulk} is the energy of an equivalent number of bulk units and A is the surface area. A low positive value for γ indicates a stable surface.

If water is present, there will be an extra term for the energy of the bulk water in equation 3.1, in addition to the energy of the bulk mineral, and E_{surf} becomes the

total energy of the hydrated surface comprising both the hydroxyapatite and the water in the simulation cell.

$$\gamma = \frac{\left(E_{hyd-surf} - E_{bulk} - E_{water}\right)}{A} \tag{3.2}$$

The adsorption energies of different molecules at a surface can be calculated by comparing the energy of the surface with the adsorbed surfactant, and the sum of the energies of the pure surface and that of an isolated molecule.

$$E_{ads} = E_{system} - (E_{surf} + E_{molecule}) \tag{3.3}$$

where E_{system} is the energy of the surface with the adsorbed molecule, E_{surf} is the energy of the simulation cell containing the surface only and $E_{molecule}$ is the energy of the molecule in its optimised geometry, calculated using the same conditions. A negative adsorption energy thus indicates that adsorption of the molecule at the surfaces is energetically favourable.

When water is present in the simulations, the strength of binding of the molecule to the surface is expressed by the adsorption energy, calculated as given in equation 3.4

$$E_{ads} = E_{system} - (E_{hyd-surf} + E_{hyd-molecule}) \tag{3.4}$$

where E_{system} is the energy of the complete system also containing the water molecules, $E_{hyd\text{-}surface}$ is the energy of the hydrated surface and $E_{hyd\text{-}molecule}$ is the energy of the solvated molecule. Again, a negative adsorption energy here indicates that adsorption of the molecule at the surfaces is favourable with respect to the surface and molecule being separated in the aqueous environment.

3.2 Computational tools

The methodology used involved a combination of two different computational techniques: electronic structure calculations and interactomic potential-based methods. The *ab initio* methods have been employed to validate the interatomic potential models used for the classical calculation, as the latter allow the simulation of larger and more complex systems. There are a number of computer programs that can be used for these two types of calculations. In this study, we have used two codes: SIESTA and DL_POLY. The basic features of each code are explained below.

3.2.1. Electronic Structure Calculations: SIESTA program

SIESTA (Spanish Initiative for Electronic Simulation with Thousands of Atoms) (Ordejon *et al.*, 1996) is a density functional theory (DFT) package, based on a flexible linear combination of atomic orbitals (LCAO) basis sets. Its main advantage is an essentially perfect order-*N* scaling algorithm, which means that the CPU time used by the calculations scale linearly with the number of atoms in

the simulation cell. It therefore allows extremely fast simulations using minimal basis sets and very accurate calculations with complete multiple zeta and polarized bases, depending on the required accuracy and available computational power. In recent years there has been a wide range of studies performed with SIESTA, including metallic surfaces (Coquet *et al.*, 2006, Fernández *et al.*, 2007), nanotubes (Hidalgo *et al.*, 2009, Kostyrko & Krompiewski, 2008), and biomolecules (Fernandez-Serra *et al.*, 2000, Austen *et al.*, 2008).

The SIESTA code is based on both the local (spin) density approximation (Perdew *et al.*, 1981) (LDA/LSD) and the generalized gradient approximation (GGA) (Perdew *et al.*, 1996). It uses standard norm-conserving pseudopotentials (Hamann *et al.*, 1979, Bachelet *et al.*, 1982) in the their fully nonlocal form (Kleinman & Bylander, 1982). Also, it has included scalar-relative effects and the nonlinear partial-core correction to treat XC in the core region (Soler *et al.*, 2002). In our calculations, the total energy and structure of the simulation system was determined with the Perdew-Burke-Ernzerhof (Perdew *et al.*, 1996) generalized gradient approximation for the exchange-correlation functional. For all the atoms, the pseudopotentials were generated in the Troullier-Martins manner (Troullier & Martins, 1991). The basis sets used were of the DZP type (double ζ with polarization), with the basis sets of both the hydroxy oxygen and molecular oxygen obtained from the optimisation of water at 0.2 GPa (Anglada *et al.*, 2002, Junquera *et al.*, 2001), while for the carbon, nitrogen and hydrogen the basis sets were optimised with respect to the interaction of guanine-cytosine at 0.2 GPa (Fernandez-Serra *et al.*, 2000).

The basis functions and the electron density are projected onto a uniform real-space grid in order to calculate the Hartree and exchange correlation potentials and matrix elements. The mesh size of the grid is controlled by an energy cutoff which sets the wavelength of the shortest plane wave that can be represented on the grid; we have taken a cutoff value of 250 Ry, and a force tolerance of 0.01 eV/Å. In order to limit the range of the basis pseudoatomic orbitals (PAO), a common energy shift is applied, here equal to 0.02 Ry for geometry optimisations and the basis functions are truncated at the resulting radial node.

To obtain the electronic structure of the system the standard diagonalisation solution of the Hamiltonian is used in preference to the order-N, linear-scaling method. For the systems we study here (less than 250 atoms), the diagonalisation method is competitive and more convenient than the order-N method. A tolerance of 5×10^{-5} is used for the density matrix to define when selfconsistency has been achieved. For geometry optimisations the conjugate-gradient approach was used with a threshold of 0.01 eV $Å^{-1}$.

3.2.2 Interactomic Potential-Based Simulation

The force fields used in atomistic modelling are primarily designed to reproduce structural properties, however, all parameters of the potential functions need to be fitted. There are two principal methods of deriving potential parameters, empirical fitting and quantum mechanics calculations. Empirical methods attempt to derive the interatomic potential parameters by fitting to known experimental data of the system, such as structure, elastic and dielectric constants, phonon dispersion

curves, and vibrational frequencies. Unfortunately, experimental data sometime may be unavailable for particular classes of systems. *Ab initio* methods can then be used to obtain the data for the parametrization, such as the equilibrium geometry and the energy surface corresponding to a particular structural distortion. Transferability of the functional form and parameters is an important feature of a force field, and should always be checked carefully if they are going to be applied to a system other than the one used in the parameterisation.

Another important issue is the combination of different interatomic potential sets. This is particularly relevant in the simulation of biomaterials and biomineralization processes, because this type of simulation requires the inclusion of both inorganic and organic molecules in solvent environments. A careful selection of interatomic cross-term potentials is necessary. For the simulations in these fields, the Lorentz-Berthelot mixing rules (Hirschfelder *et al.*, 1954) are extensively used to calculate the intermolecular potential between pairs of non-identical atoms (Harding & Duffy, 2006, Ravichandran *et al.*, 2001, Gao & Kong, 2004, Chen *et al.*, 2007).

$$r_{ij}^{*} = \frac{\left(r_{ii}^{*} - r_{jj}^{*}\right)}{2} \tag{3.5}$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \tag{3.6}$$

where ε_{ii} and ε_{jj} are the well depths of potential for atoms with type *i* and type *j* in the parameter of Lennard-Jones potential, r_{ii}^{*} and r_{jj}^{*} are the distances at which the well depth occurs for atoms with type *i* and *j*. r_{ij}^{*} and ε_{ij} are the parameters to

calculate the cross interaction between *i* and *j* atoms. However, previous studies (Duffy & Harding, 2004a, Cormack *et al.*, 2004) have demonstrated that potentials obtained in this way severely over-estimate binding energies at interfaces.

An alternative approach has been used by Schröder *et al*. (Schroder *et al.*, 1992) to investigate the interaction of zeolites with methanoic acid. This method recasts the interatomic cross-term potentials to fit the different charges between the systems and can be equally applied to biominerals, as has been demonstrated by de Leeuw *et al*. (de Leeuw *et al.*, 1998) and Duffy and Harding (Duffy & Harding, 2004a) with success. Freeman *et al* (Freeman *et al.*, 2007) have developed a methodology that atoms are assigned charges that maintain the ratio of charges between different atoms but also mimic the electrostatic interactions between organic and mineral components, i.e., a reduced charge model.

A number of interactomic potential models are available for modelling organic macromolecules. They include the CVFF (Brown & Chow, 1976), DREIDING (Mayo *et al.*, 1990), and AMBER (Cornell *et al.*, 1995) force fields, which use relatively small charges on the atoms together with strong bond terms. In many cases, force fields describing inorganic materials are based on the Born model of solids (Born & Huang, 1954), which assumes that the ions in the crystal interact via long-range electrostatic forces and short-range forces, including both the repulsions and van der Waals attractions between neighbouring electron charge clouds, which are described by simple analytical functions.

3.2.2.1 Potential parameters of Hydroxyapatite

We have used the potential model developed by De Leeuw *et al* for the apatite mineral (de Leeuw, 2004b), which includes electronic polarisability via the shell model of Dick and Overhauser (Dick & Overhauser, 1958). Within this model, each polarisable ion (oxygen in our case) is represented by a core and a massless shell, connected by a spring. The polarisability of the oxygen ions is then determined by the spring constant and the charge of the shell. When necessary, angle-dependent forces can also be included to allow directionality of bonding, for example in the covalent phosphate anion (Table 3.1).

Table 3.1 Hydroxyapatite Potential

Ion	Charges (e)		Core-shell interaction kcal Å^2
	Core	Shell	
Hydroxyapatite			
Ca	+2.000		
P	+1.180		
Phosphate oxygen (O2)	+0.587	-1.632	11693.0837
Hydroxy oxygen (O1)	+0.900	-2.300	1726.5476
Hydroxy hydrogen (H1)	+0.400		
	Buckingham potential		
Ion pair	A (kcal)	ρ (Å)	C (kcal Å^6)
Ca—O2 (shell)	35719.9050	0.2970	0.0
Ca—O1 (shell)	28806.3750	0.3437	0.0
O2 (shell)—O2 (shell)	377294.3772	0.2130	79.9665
O1 (shell)—O1 (shell)	524598.6564	0.1490	160.6243
O2 (shell)—O1 (shell)	524598.6564	0.1490	113.3819
H1—O2 (shell)	7190.0712	0.2500	0.0

H1—O1 (shell)	7190.0712	0.2500	0.0
	Morse potential		
	D (kcal)	α ($Å^{-1}$)	r_0 (Å)
P—O2 (core)	79.9665	2.030	1.600
H1—O1 (core)	162.5256	3.1749	0.9485
	Three-body potential		
	K (kcal rad^{-2})		θ_0
O2(core)—P—O2(core)	30.4800		109.47

3.2.2.2 Potential parameters of amino acids

The AMBER force field parameters were used to model the amino acid molecules in our study. In this force field, bond and angles are represented by a simple diagonal harmonic expression, the VdW interaction is represented by a Lennard-Jones (6-12) potential, electrostatic interactions are modeled by a Coulombic interaction of atom-centred point charges, and dihedral energies are represented with a simple set of parameters, often only specified by the two central atoms. Electrostatic and van der Waals interactions are only calculated between atoms in different molecules or for atoms in the same molecule separated by at least three bonds.

For the glycine amino acid molecule (—NH-CH_2-COO^- or —NH-CH_2-CO—) the atoms are labelled as N for the amino or amine atom; CA for the α-carbon atom; C for the carbonyl or carboxylate carbon atom; O for the carbonyl or carboxylate oxygen; Ha for the hydrogen atom attached to the α-carbon atom, and H for the hydrogen atom attached to the amino group. For the proline amino acid residue (—NH-(CH_2-CH_2-CH_2)-CH-CO—) the atoms are labelled as N for the amino or

amine atom; CA for the α-carbon atom; C for the carbonyl or carboxylate carbon atom; O for the carbonyl or carboxylate oxygen; CB, CG, and CD for the carbon ring atoms; the Ha for the hydrogen atom attached to the α-carbon atom, and Hb, Hg and Hd for the hydrogen atom attached to the carbon ring atoms. In case of hydroxyproline (—NH-(CH_2-CH(OH)-CH_2)-CH-CO— or $^{+}H_2N$-(CH_2-CH(OH)-CH_2)-CH-CO—), there are two atoms more, calling OD the hydroxyl oxygen atom and HO the hydroxyl hydrogen (see Table 3.2).

Table 3.2 Charges of the amino acids from AMBER force field

Atom	Charges (e)	
GLY	Residue -NH-CH_2-CO-	C-terminal -NH-CH_2-COO^-
N	-0.5157	-0.3821
CA	-0.0252	-0.2493
C	+0.5973	+0.7231
O	-0.5679	-0.7855
Ha	+0.0698	+0.1056
H		
PRO	Residue -N-(CH_2-CH_2-CH_2)-CH-CO-	
N	-0.2548	
CA	-0.0266	
C	+0.5896	
O	-0.5748	
CB	-0.0070	
CG	+0.0189	
CD	+0.0192	
Ha	+0.0641	
Hb	+0.0253	
Hg	+0.0213	
Hd	+0.0391	
HYP	Residue -N-(CH_2-CH(OH)-CH_2)-CH-CO-	N-terminal H_2N-(CH_2-CH(OH)-CH_2)-CH-CO-
N	-0.2548	-0.2548
CA	+0.0047	+0.0047
C	+0.5896	+0.5896
O	-0.5748	-0.5748
CB	-0.0203	-0.0032

CG	+0.0400	+0.0908
CD	+0.0595	-0.0081
OD	-0.6134	-0.5930
Ha	+0.0770	+0.1135
Hb	+0.0426	+0.1135
Hg	+0.0416	+0.1135
Hd	+0.0070	+0.1135
HO	+0.3851	+0.4588
H		+0.3045

3.2.2.3 Potential parameters of water molecules

The water model used was TIP3P (Walser *et al.*, 2001), which have dipole moments that are about 20% higher than the gas-phase value for water. This water model, which has empirically derived charges, includes condensed-phase electronic polarization implicitly.

Table 3.3 Charges of water molecules

Atom	Charges (e)
	Core
Water	
Oxygen (OW)	-0.8340
Hydrogen (HW)	+0.4170

3.2.2.4 Cross-interactions between HA, the organic molecules and water

As described above, inorganic materials contain relatively large partial charges distributed over atomic sites, whereas biomolecules force fields typically use partial charges of relatively smaller value. In such a case, application of the Lorentz-Berthelot mixing rules for nonbonded interactions gives problematic behaviour. However, in our case as the charges from the De Leeuw and AMBER potentials are similar, the partial charges on the hydroxyapatite are in line with the partial charges on water and biomolecules. Therefore, we have made use of the Lorentz-Berthelot mixing rules for the biomolecules/hydroxyapatite/water

interactions. The parameters for the interaction of calcium with the peptide, which were added to the AMBER force field, were first applied by Marques and co-workers (Robinson *et al.*, 2006).

The reliability of the methodology used has been tested by comparing with *ab initio* calculations. In particular, we have used this procedure to calculate the adsorption energies for peptide molecule and the hydroxyapatite surfaces, and we found that the results agree well with DFT calculation of the same systems in both the mode and strength of binding of the peptide to the surface. Also, our tests for the interactions between hydroxyapatite and water molecules showed good agreement with previous results, using other potentials.

3.2.5 DL_POLY program

DL_POLY code (Todorov *et al.*, 2006) applies methods based on interatomic potential; designed to facilitate Molecular Dynamics (MD) simulations of macromolecules, polymers, ionic systems and solutions on a distributed memory parallel computer. We have used this program to incorporate the effect of the temperature and solvation in our calculations. This program was written to allow almost any functional form of interatomic potential to be used in the simulations. For the two body terms, the standard harmonic term is available as well as the Lennard-Jones, Buckingham, Morse, and more complex potential functions. Similarly for three-body terms, there are harmonic, harmonic + exponential, harmonic cosine, and others available. All common forms of non-bonded atom-

atom potential can be used. It is also relatively easy to adapt force field from other libraries to DL_POLY.

In MD simulations, there are two ways of treating the shells, which are essentially massless and therefore do not obey the normal dynamic laws. One is to perform a shell-only energy minimisation at each time step (Lindan & Gillan, 1993), and the other is the approach used in DL_POLY, where a small mass is assigned to the shells (Vocadlo *et al.*, 1997, Ferneyhough *et al.*, 1994). In this study, we have chosen a mass of 0.2 a.u. for the oxygen shell, which is small enough compared to the mass of the hydrogen atom (1.0 a.u.), in order to avoid energy exchange between the vibrations of oxygen core and shell and the oxygen and hydrogen vibrations. Due to the use of a shell model for oxygen, we need to use a small time step of 0.1 fs to maintain the stability of the simulation.

As described in section 2.7.3, the integration algorithms in DL_POLY are based on the Verlet leap-frog scheme (Verlet, 1967). We have used the Nosé-Hoover (Nose, 1984, Hoover, 1985) algorithm for the thermostat, as this algorithm generates trajectories in both NVT and NPT ensembles, thus keeping our simulations consistent. The Nosé-Hoover parameters were set at 0.1 ps for the thermostat and 0.5 ps for the barostat relaxation times. The simulations were initially performed using NVE and NVT ensembles to equilibrate the system for 100 ps, followed by both equilibration and the final data collection runs at NPT for at least 1 ns. The production run produces a range of properties of the system at each time step, including structure, temperature, volume and energy, as well as

the averaged properties and root mean square deviations. The energies reported in this thesis are the averages obtained at the end of each production run, whereas the structures shown in the Figures are usually snapshots, unless stated otherwise.

Chapter 4 Hydroxyapatite

In this chapter, we first investigate the geometry and relative energies of the bulk and (0001) and $(01\bar{1}0)$ surfaces of hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$), using density functional theory. Subsequently, using methods based on interatomic potentials, we extend the study to investigate the effect of temperature and hydration on the surface stabilities.

4.1 Introduction

Calcium phosphate apatites or apatites have the general formula $Ca_{10}(PO_4)_6(X)_2$, where X is an electronegative atom or group, such as OH^-, Cl^-, F^-, or Br^- (KAY, 1964). Apatites are especially important as materials for bone-tissue engineering, due to their similar composition and biocompatibility, allowing strong interaction at the bone interface and hence giving rise to their extensive use in dentistry and medicine. In addition, these materials are also widely used as catalysts, adsorbent, and to sequester heavy metal ions, a result of both the bulk and surface properties of the compounds (Mostafa & Brown, 2006). Recently, novel apatite-type silicates have been proposed as alternative solid electrolyte materials following the exciting discovery of fast oxide-ion conductivity and potential used in fuel cells, oxygen sensors, and ceramic membranes (Tolchard *et al.*, 2007, Jones *et al.*, 2008, Nakayama *et al.*, 1995).

One of the most important apatites is hydroxyapatite (HA), $Ca_{10}(PO_4)_6(OH)_2$, which has received extensive attention by a number of research groups. Due to the

fundamental function carried out by HA, the knowledge of its structure at the atomic level is considered of great importance. Several experimental studies have analyzed samples of the pure material (stoichiometric and non-stoichiometric) and of the material after immersion in simulated body fluid, using a range of techniques including X-ray diffraction (Murugan & Ramakrishna, 2005), IR and Raman spectrometry (Rehman & Bonfield, 1997, Tsuda & Arends, 1994, Cusco *et al.*, 1998), microcalorimetry (Cabanas *et al.*, 2002, Rautaray *et al.*, 2005), NMR (Qiu *et al.*, 2005), and many others. Several computational studies of HA have also been conducted, both with molecular mechanics and by means of electronic structure methods (Calderin *et al.*, 2003, Haverty *et al.*, 2005, de Leeuw, 2002, Corno *et al.*, 2006), as in this studies.

The crystal structure of hydroxyapatite has the $P6_3/m$ space group with the 4e Wyckoff positions occupied by two hydroxy oxygen atoms, each with 1/2 occupancy (Stork *et al.*, 2005). In order to translate this structure into a model with full occupancies for the hydroxy groups, as is required to carry out the calculations, we have assigned alternate 0 and 1 occupancies to these sites along the hydroxy channels in the *c* direction, thus changing the space group of the hydroxyapatite unit cell from $P6_3/m$ to $P6_3$. Since there is only one hydroxy channel per unit cell, all the hydroxy groups in the periodic structure are oriented in the same direction, creating a net electric polarisation per unit cell. In the real structure, disorder in the relative orientation of the parallel OH channels means that electric polarisation is not present in the material. Therefore, in order to create a more realistic structure, we have doubled the cell in the *b* direction, and assigned opposite orientations to the two OH channels in this supercell. This antiparallel

orientation of the hydroxy groups is calculated to be 19 kJ mol^{-1} more stable than the parallel configuration, in agreement with previous theoretical results (de Leeuw, 2001, 2002).

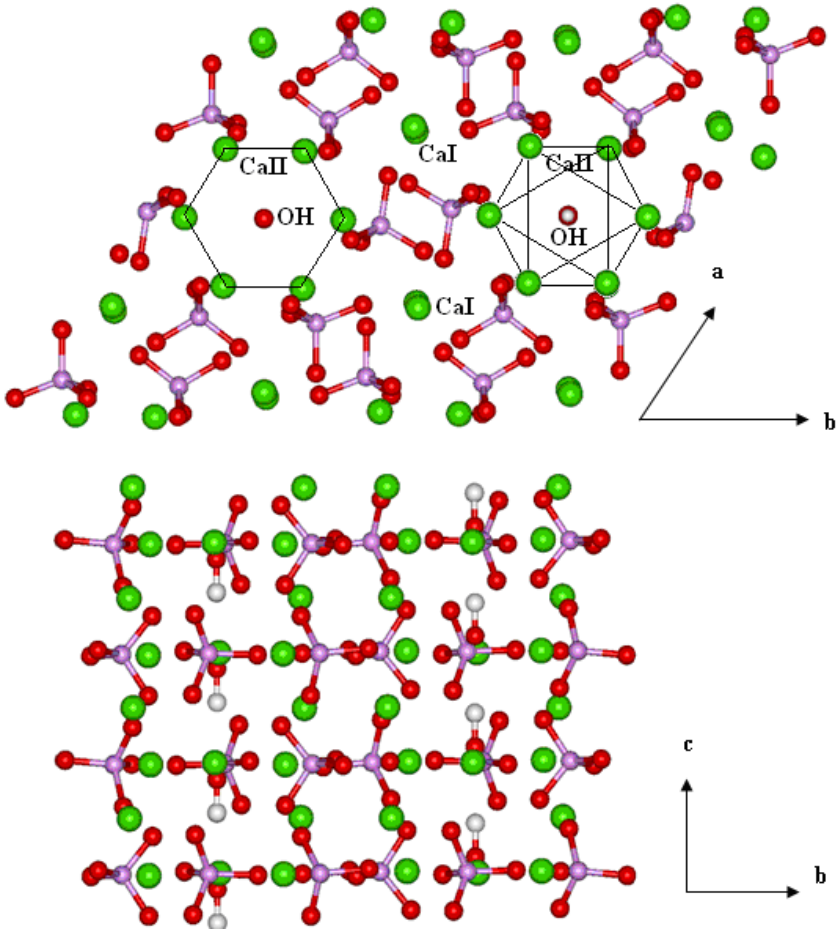


Figure 4.1 Plan and side views of the hydroxy apatite structure, showing the OH groups in hexagonal channels surrounded by triangles of Ca ions, (Ca = green, O = red, P = purple, H = white).

The configuration with antiparallel hydroxy groups in neighbouring columns, shown in Figure 4.1, coincides with the arrangement of the hydroxy groups in pure, synthetic HA, which crystallises as an ordered monoclinic structure ($P2_1$) with a double unit cell compared to the hexagonal structure (Stork *et al.*, 2005). However, as natural HA has the hexagonal structure, we have used this structure

for our calculations, thus allowing direct comparison of the simulations with experimentally determined surfaces.

4.2 DFT Calculations of the bulk Hydroxyapatite crystal

The simulations were performed using the SIESTA program (Ordejon *et al.*, 1996), details of the simulation were describe in section 3.2.1. The suitabilities of cutoff energy, *k* points and the force tolerance were evaluated by monitoring the convergence of the total energy with respect to the various parameters and validate against the structural properties of the hydroxyapatite. Convergence in the energy and lattice parameters is achieved with a 2x1x3, net as shown in Table 4.1.

Table 4.1 Convergence of lattice parameters and energy with k-point sampling

k-points	a(Å)	2b(Å)	c(Å)	Energy (eV)
1x1x1	9.40	18.81	6.95	-44980.31
2x1x3	**9.40**	**18.81**	**6.95**	**-44981.40**
4x2x6	9.41	18.82	6.93	-44981.40

The cell parameters of the relaxed structure, shown in table 4.1, are in good agreement with experiment ($a = b = 9.43$ Å and $c = 6.88$ Å (Posner *et al.*, 1958), discrepancy <1%), as well as with previous theoretical results (de Leeuw, 2002, Ellis *et al.*, 2006).

The calcium and phosphate ions respond little to the location of the hydroxy groups, which, as described, are stacked in regular columns within the hexagonal channels formed by triangles of Ca ions, although the direction of the OH groups may differ randomly between neighbouring channels without a significant energetic penalty (de Leeuw, 2001, 2002). The PO_4 groups display little variation in size or shape, and the calculations yield an average P-O bond length of 1.59 Å and calculated average O-P-O angle of 109.41°, compared to the corresponding experimental averages of 1.54 Å and 109.45°, respectively.

In the reduced symmetry, there is only one type of Ca(II) position. However, there are two types of Ca(I) in the cell, which differ slightly in the distances to the neighbouring oxygen atoms and are labelled as Ca(Ia) and Ca(Ib), in which Ca(Ib) has three oxygen atoms as first neighbours, Ca(Ia) has six (Corno *et al.*, 2006, Ellis *et al.*, 2006). In one unit cell (44 atoms), there are two ions of type Ca(Ia) and two of Ca(Ib) involved in an octahedral arrangement and, for phosphates groups, there are six of them inside, all equivalent for symmetry. Table 4.2 shows the average values of Ca(I)···O co-ordination distances, compared to the experimental values. The Ca(Ib) site is thus distinguished from the Ca(Ia) site by increased distances to the third neighbour atoms.

Table 4.2 Co-ordination distances around Ca (Ia) and Ca (Ib).

Ca site	Exp.	Calc.	
	Ca (I)	Ca (Ia)	Ca (Ib)
Ca(I)-O1	2.42	2.40	2.34
Ca(I)-O1	2.42	2.41	2.38
Ca(I)-O1	2.42	2.42	2.39
Ca(I)-O2	2.45	2.44	2.45
Ca(I)-O2	2.45	2.45	2.46
Ca(I)-O2	2.45	2.45	2.52
Ca(I)-O3	2.77	2.59	2.91
Ca(I)-O3	2.77	2.62	2.90
Ca(I)-O3	2.77	2.69	3.07

Turning now to the Ca(II) site, the OH group is linked to three Ca(II) ions, which form and equilateral triangle in the *ab* plane with the hydroxyl in the center (see Figure 4.1). The calculated Ca(II)···O co-ordination distances are listed in table 4.3, which compare favourably with the experimental values.

Table 4.3 Co-ordination distances around Ca(II).

Ca site	Exp.	Calc.
	Ca (II)	Ca (II)
CaII)-O1	2.68	2.62
Ca(II)-O2	2.35	2.35
Ca(II)-O3	2.34	2.34
Ca(II)-O3	2.34	2.34
Ca(II)-O3	2.50	2.40
Ca(II)-O3	2.50	2.55
Ca(II)-OH	2.38	2.38

Geometry optimization gives bond distances and angles in good agreement with experimental values, as well as the previous theoretical result of the Leeuw *et al*.(de Leeuw, 2004b), and Corno *et al*.(Corno *et al.*, 2006). As already discussed, the most plausible physical interpretation of the OH disorder is the monoclinic arrangement, and, in relation to the perturbation of the PO_4 tetrahedra, the hexagonal represents the average of the $P2_1$ positions. Based on neutron diffraction data, it has previously been suggested (Sanger & Kuhs, 1992) that $P6_3/m$ is not the true symmetry of HA and represents only the average structure on a macroscopic level.

4.3 Hydroxyapatite surfaces: DFT calculations

We created a range of surfaces with low Miller indices, as these surfaces have the largest inter-planar spacings and as a result are generally the most stable, which were subsequently optimised using the SIESTA method (Ordejon *et al.*, 1996). In this study, we have concentrated on two of the most significant surfaces of the apatite bone mineral, namely the (0001) and $(01\bar{1}0)$ surfaces. These two surfaces are important in the morphology of the apatite mineral platelets and there is experimental evidence that these faces provide binding sites for many ionic species, small molecules, polymers and anionically-modified cell surfaces (Wierzbicki & Cheung, 2000, de Leeuw, 2004b).

The basic building block of the (0001) surface consists of three phosphate groups and three calcium ions, which is translated along the *a* and *b* axes to form the

infinite surface. Previous energy minimisation calculations of the different terminations of a number of HA surfaces (Filgueiras *et al.*, 2006) have shown that this (0001) termination is the most stable, hence its use in this work. It is not possible to construct a HA slab of the $(01\bar{1}0)$ surface with the same surface termination on both sides, even though the total dipole moment across the slab perpendicular to the surface is zero, which is the requisite for a stable surface. As a result the bottom plane of the slab contains two calcium ions only whereas the top plane has two extra hydroxy groups. The (Ca-Ca) termination is the more stable, but the (OH-OH-Ca-Ca) has a greater variation of surface species, which could form different interactions with the adsorbates (Filgueiras *et al.*, 2006), and both planes will occur during the HA growth process. The surface energy of this slab will be an average of the two $(01\bar{1}0)$ terminations, but the adsorption energy will be specific to the particular plane used as substrate. For our adsorption studies, we have used the more reactive (OH-OH-Ca-Ca) termination, as this provides a wider variety of adsorption sites and hence will provide more insight into the different modes of adsorption.

The HA surfaces were modelled as slabs of 10 Å thick, perpendicular to the surface, and infinitely repeated laterally through 3-dimensional periodic boundary conditions. The slabs were separated from their images in the neighbouring cells by a vacuum region of approximately 50 Å, thus avoiding interactions between repeating slabs. Tests of the convergence of the total energy and properties with respect to the slab thickness showed that the surface energies are converged within 0.05 Jm^{-2} (see Table 4.4), whereas on increasing the slab thickness the

adsorption energies and indeed the structures of the surfaces were not affected significantly (within 1 $kJmol^{-1}$), as described in chapter 5.

The calculated surface energies, listed in Table 4.4 indicate that the (0001) surface is more stable than the (average) $(01\bar{1}0)$ surface, in agreement with other *ab* initio and classical calculations (Filgueiras *et al.*, 2006, de Leeuw & Rabone, 2007, Corno *et al.*, 2009). The less stable surface will be more reactive and may quickly accumulate more HA material to eventually grow out of the crystal morphology. However its higher reactivity may also lead to a stronger interaction with the collagen, which is important for the hydroxyapatite formation at the organic matrix.

Table 4.4 Surface energies of the surface simulation cell

Surface Energy γ / Jm^{-2}				
Surface	Termination	2x2x1 (176 Atoms)	2x2x2 (264 Atoms)	2x2x3 (528 Atoms)
(0001)	PO_4- Ca- PO_4- PO_4-Ca-Ca	1.01	0.98	0.94
$(01\bar{1}0)$	OH-OH-Ca-Ca Ca-Ca	1.32	1.30	1.27

The relaxed (0001) and $(01\bar{1}0)$ surfaces are shown in Figures 4.2 (a) and (b), respectively. After geometry optimisation, the symmetry is broken due to the re-orientation of the hydroxy groups within the channel in both surfaces. Upon geometry optimisation of the bulk crystal, the hydroxy groups remained lined up within the columns, where the OH···OH hydrogen bond distance averages 2.49 Å, in agreement with previous DFT calculations using a plane-wave pseudo-potential

approach (de Leeuw, 2001). However, when a surface is present, the now under-coordinated surface hydroxy groups rotate to increase their interaction with each other and with phosphate oxygen atoms, as shown in Figure 4.2.

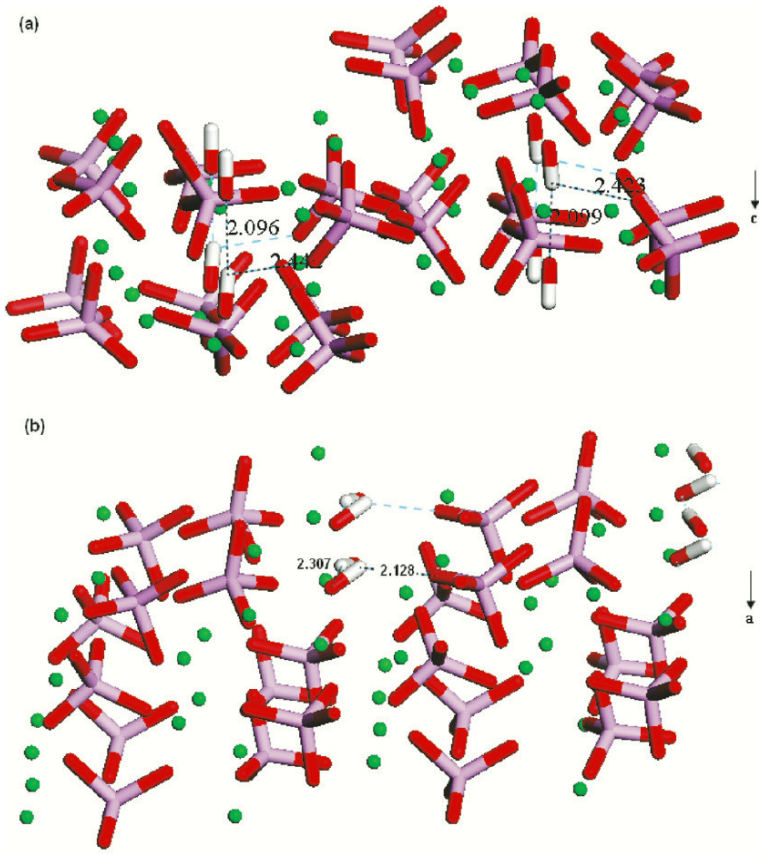


Figure 4.2 Relaxed surfaces of hydroxyapatite: **(a)** PO_4-Ca-PO_4-PO_4-Ca-Ca termination of the (0001) surface and **(b)** Ca-Ca in the bottom and (OH-OH-Ca-Ca) in the top of the $(01\bar{1}0)$ surface (Ca = green, O = red, P = purple, H = white).

In the (0001) surface, we observe that the surface calcium ions have relaxed further into the surface, decreasing distances to oxygens of the phosphate groups by 6%, where the Ca···O bond length to the oxygen of the phosphate groups is now 2.43 Å (2.45 Å average value in the bulk). The column of hydroxy groups has become distorted with the topmost group relaxing out of the surface, whereas

the OH···OH hydrogen-bond distances have shortened to an average value of 2.09 Å and hydrogen-bonds between OH···OPO_3 are formed at about 2.44 Å.

In the $(01\bar{1}0)$ surface, the disordering of hydroxy groups is more significant than in the (0001) surface, probably due to their initial orientation parallel to the surface plane, although the surface area of the (0001) is greater than that of the $(01\bar{1}0)$ surface and will therefore allow more surface relaxation, if required. In the (0001) surface, the surface hydroxy groups remain in their original orientation, but they relax into the bulk material. In the $(01\bar{1}0)$ surface, the hydroxy groups remain more or less parallel to the surface plane, although half of the groups rotate to interact more closely with neighbouring OH groups with average OH···OH distances of 2.31 Å, whereas the resulting OH···OPO_3 distances are about 2.13 Å. The surface calcium ions interact with the oxygen atoms of the hydroxy groups, at a Ca···OH^- bond length of 2.2 Å. There is no significant movement of the surface calcium ions or sub-surface phosphate groups.

4.4 Hydroxyapatite: MD calculations

We were not only interested in the hydroxyapatite bulk and surface structure but also in the effect of water on the structure and stabilities of the hydroxyapatite surfaces. Therefore we next employed atomistic simulations, which are better suited to investigate the necessarily larger simulation cells, to study a series of dry and hydrated hydroxyapatite surfaces. However, in order to validate the potential

model, we first compared the results from the DFT calculations with those using atomistic simulations, and hence we once again simulated the bulk materials and surfaces.

We have used the de Leeuw potentials to calculate the cell parameters for the hydroxyapatite bulk, which are list in Table 4.5 and show good agreement with the experimental values (Posner *et al.*, 1958) and other force fields (Lee *et al.*, 2000, Hauptmann *et al.*, 2003). The good agreement between the DFT and classical MD simulations simulation is also encouraging as these results were obtained independently and the potentials were not fitted to the DFT results. We next extended our study to atomistic simulations of the two surfaces and their interaction with water.

Table 4.5 Lattice parameters of hydroxyapatite crystal

Parameter	Exp	DFT	**Leeuw *et al.***	Lee *et al*.	Hauptmann *et al*.
a (Å)	9.432	9.40	**9.38**	9.53	9.46
b (Å)	9.432	9.41 (2b=18.81)	**9.38 (2b=18.75)**	9.53	9.46
c (Å)	6.881	6.95	**6.85**	6.61	6.90
α (°)	90	90	**90**	90	90
β (°)	90	90	**90**	90	90
γ (°)	120	119.9	**120**	120	120

4.4.1 Dry surfaces

Having optimized the hydroxyapatite bulk structure, we have created the dry (0001) and $(01\bar{1}0)$ surfaces. Upon simulation of the (0001) surface at $0K$ (shown in Figure 4.3), we observe considerable relaxation of the first and to a lesser extent second and third surface layers. The phosphate groups remain intact, with just a little rotation of the oxygen in the topmost surface layer towards the surface calcium ion, although the O-P bond length remains 1.60 Å. The Ca···O bond lengths to the oxygen of the phosphate groups below also remain approximately the same at 2.42 Å, but the calcium ion has also relaxed further into the surface to form three additional bonds to the rotated surface oxygen of the phosphate groups, with a short bond length of 2.32 Å. The column of hydroxy groups has became distorted with the topmost group relaxing out of the surface, while the second group moves more into the bulk. We label the OH as OH1-S, OH2-subS, OH1-I, OH2-I, OH1-NB and OH2-NB, where S, subs, I, and B stand for "surface," "subsurface," "intermediate," and "near bulk," respectively and the first numeral denotes the column (see Figure 4.3). The resulting OH1-S···OH1-I and OH1-I···OH1-NB distances become 2.60 and 2.54 Å, respectively, i.e. considerably longer than in the bulk material (2.39 Å). In contrast, the OH2-subS···OH2-I and OH2-I···OH2-NB distances have shortened to 2.36 and 2.02 Å. In the topmost surface layer the hydrogen-bonds between OH···OPO_3 groups are formed at about 2.44 Å.

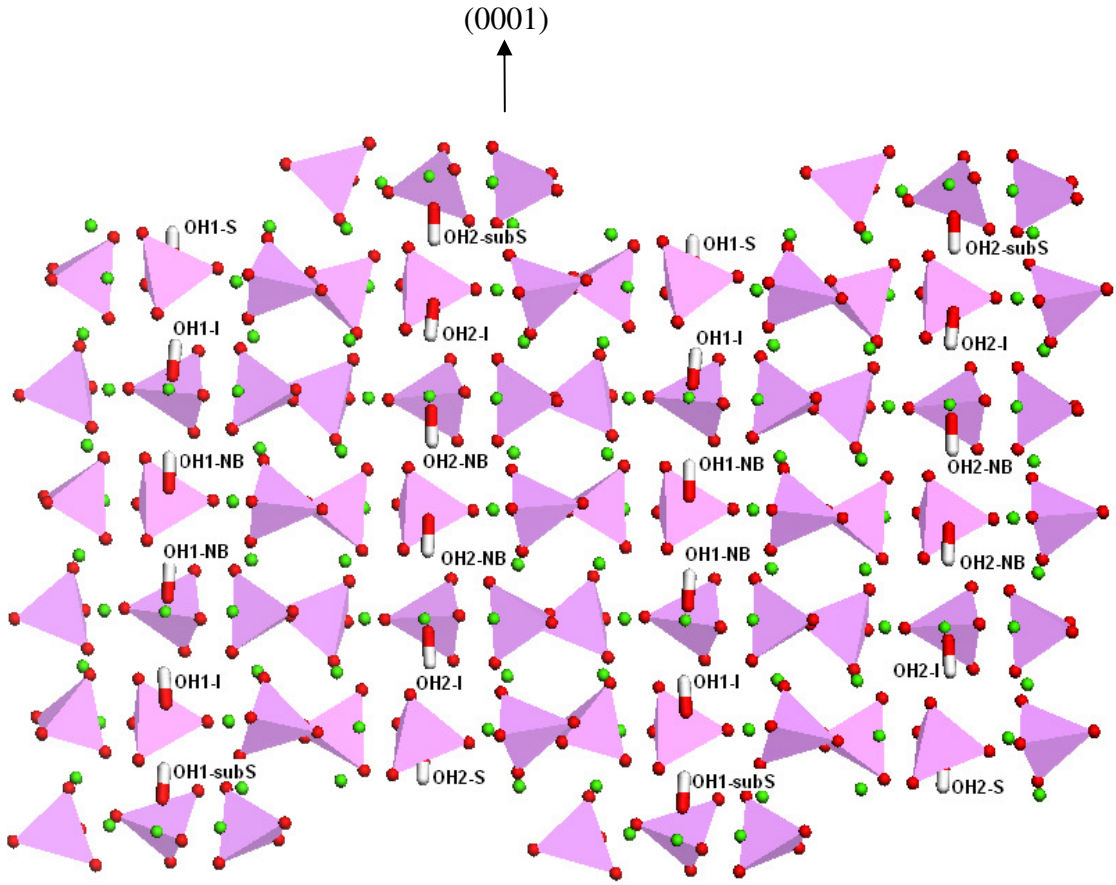


Figure 4.3 Side view of the geometry optimised (0001) surface from the MD calculation (Ca = green, O = red, P = purple, H = white, PO_4 groups displayed as tetrahedra).

In the $(01\bar{1}0)$ surface, the columns of hydroxy groups lie in the plane of the surface, rather than perpendicular to the surface as in the (0001) surface. Upon geometry optimisation (Figure 4.4) the calcium ions in the surface form new bonds with surface oxygen atoms (2.27 Å), which have become accessible through the rotation of the surface phosphate groups. The hydroxy groups remain more or less parallel to the surface plane, although half of the groups rotate to interact more closely with neighbouring OH groups with average OH···OH distances of 2.31 Å, whereas the resulting OH···OPO_3 distances are about 2.13 Å. The surface calcium ions interact with the oxygen atoms of the hydroxy groups, at

a Ca···OH distance of 2.23 Å. There is no significant movement of the sub-surface calcium ions and phosphate groups.

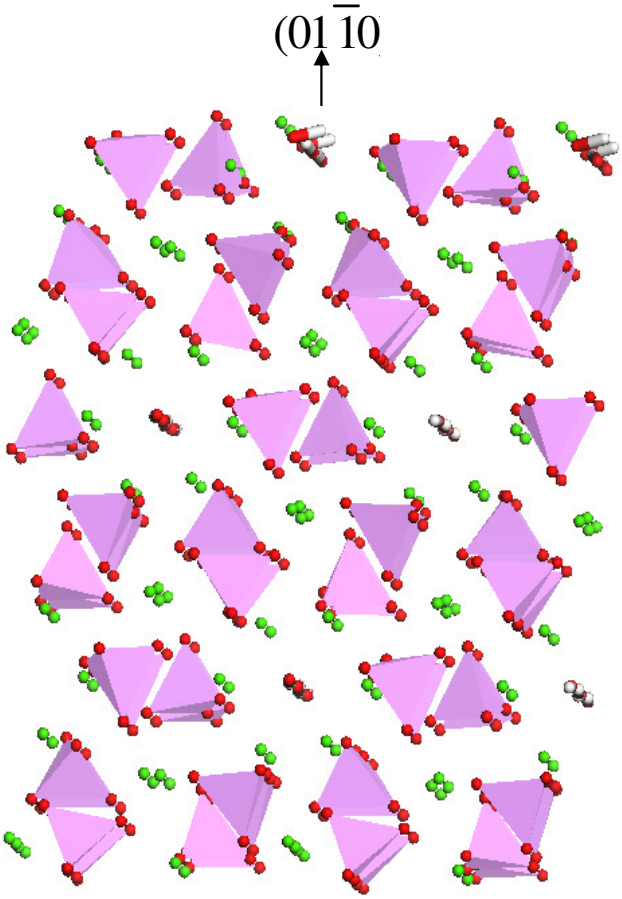


Figure 4.4 Side view of the geometry optimised $(01\bar{1}0)$ surface from the MD calculation (Ca = green, O = red, P = purple, H = white, PO_4 groups displayed as tetrahedra).

The calculated surface energies of the (0001) and $(01\bar{1}0)$ hydroxyapatite surfaces, using both DFT and MD methods, are listed in Table 4.6, which show that the (0001) surface is more stable than the average of the two $(01\bar{1}0)$ terminations. It is clear from Table 4.6 that the surface energies obtained from the two methods agree remarkably well, giving us confidence that the interatomic potential describes the hydroxyapatite surfaces accurately.

Table 4.6 Surface area *A* and surface energy γ for HA surfaces, using DFT and classical MD methods.

HA surface	A(Å^2)	γ (Jm^{-2})		
		DFT in *vacuo* 2x2x3	MD (0*K*) in *vacuo* 2x2x3	MD (310*K*) in water 2x2x3
(0001)	353.8	0.94	0.91	0.37
$(01\bar{1}0)$	261.6	1.27	1.23	0.90

4.4.2 Hydrated surfaces

The hydrated (0001) surface is shown in Figure 4.5. The water molecules do not adsorb in a regular patter, which is sometimes found on highly ordered mineral surfaces, for example the dominant calcite and α-quartz surfaces (Kerisit *et al.*, 2005, Du & de Leeuw, 2006). The HA (0001) with its surface calcium vacancies is a much less regular surface than these calcite and α-quartz surfaces, and instead of an ordered water layer, the water molecules cluster around the remaining surface calcium ions and even penetrate into the channels of the hydroxy groups. The major interaction of the water molecules with the surface is to the surface calcium ions at distances of between 2.26 and 2.63 Å.

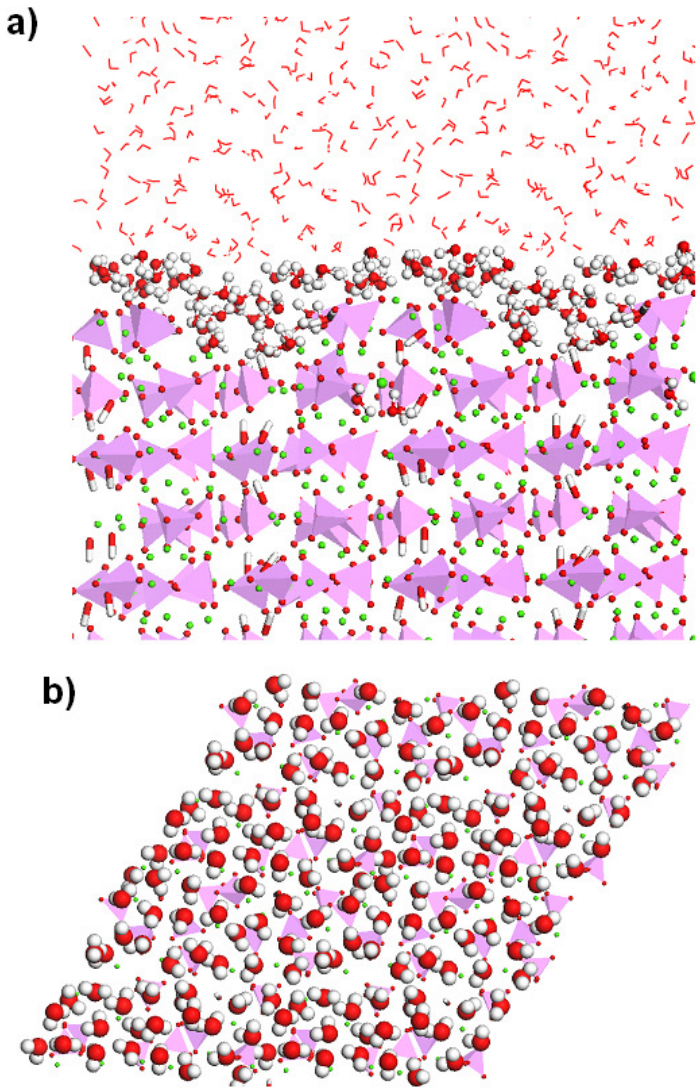


Figure 4.5 (a) Side view of the structure of interfacial water molecules near the (PO_4-Ca-PO_4-PO_4-Ca-Ca) termination of the (0001) surface and **(b)** plan view, showing an irregular adsorption pattern (Ca = green, O = red, P = purple, H = white, PO_4 groups displayed as tetrahedra, water shown as small red/white ball and stick model).

Figure 4.6 shows side and plan views of the $(01\bar{1}0)$ hydrated surface, which gave similar results to the (0001) surface. Again the water molecules interact mainly with the surface calcium ions, but in this case the surface hydroxy groups dissolve into the solvent (Ca···OH = 2.70 Å), similar to previous MD simulations using a different water potential (de Leeuw, 2004a). Water molecules replace the surface

hydroxy groups in the surface channel, although the rest of the hydroxy sub-lattice remains intact, showing just a little rotation (Ca···OH = 2.49 Å, OH···OPO_3 = 2.77 Å, and OH···OH = 2.48 Å). Additional stabilisation of the water is provided by a network of hydrogen bonding between the water molecules and surface oxygen atoms (H_{water}···OPO_3 = 2.32-2.63 Å).

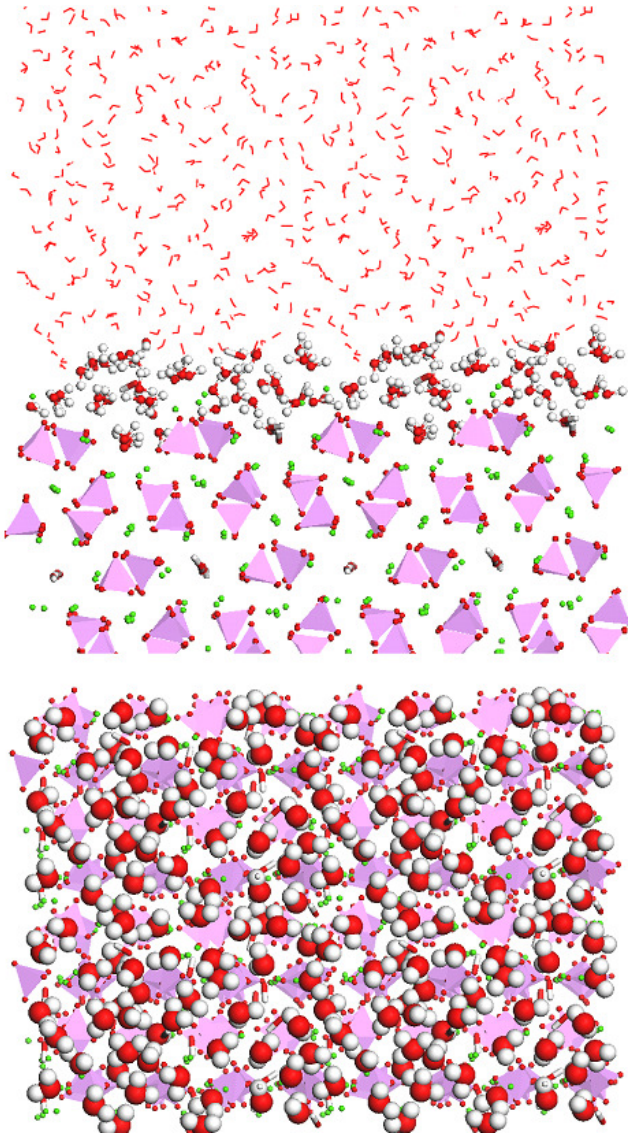


Figure 4.6 (a) Side view of the structure of interfacial water molecules near the (OH-OH-Ca-Ca) termination of the $(01\bar{1}0)$ surface and **(b)** plan view, showing an irregular adsorption pattern. (Ca = green, O = red, P = purple, H = white, PO_4 groups displayed as tetrahedra, water shown as small red/white ball and stick model).

In Figure 4.7 (a) we have plotted a diagram showing the distance between the surface, which we have defined here as the coordinate of the topmost surface Ca ions and the oxygen atom of each water molecule. Figure 4.7 (a) clearly shows that the water molecules form a well-defined monolayer, with a gap between it and the next water layer but this layering effect is lost completely at distances further away from the surface than about 4 Å. The distance from the surface of the first water layer is in agreement with experiment, which found that the height of the first water layer in hydrated fluorapatite is 2.64 Å (Park *et al.*, 2004), as already predicted in static lattice calculations of hydrated fluorapatite surfaces (Mkhonto & de Leeuw, 2002). The presence of many surface species: calcium ions, phosphate oxygens and hydroxy groups, and their surface geometry do not encourage the formation of intermolecular interactions between the adsorbed water molecules to form a regular network. This lack of structure in the adsorbed water layer leads to the bulk-like water phase at the hydroxyapatite interface, rather than the ice-like layered geometry found at the α-quartz surface (Du & de Leeuw, 2006).

The radial distribution functions (RDF) plotted for the distances between calcium ions and the oxygen atoms of the water molecules, between surface oxygen atoms of the phosphate groups and hydrogen atoms of the water molecule, and between surface hydrogen atoms of the hydroxy groups and oxygen atoms of the water molecules, are shown in Fig. 4.7 (b), where the first peaks of each at 2.38, 1.85 and 2.86 Å, respectively, are in accord with *ab initio* MD simulations of solvated Ca ions, and phosphate and hydroxy groups.(Di Tommaso & de Leeuw, 2008, Tang *et al.*, 2009a, Chen *et al.*, 2002)

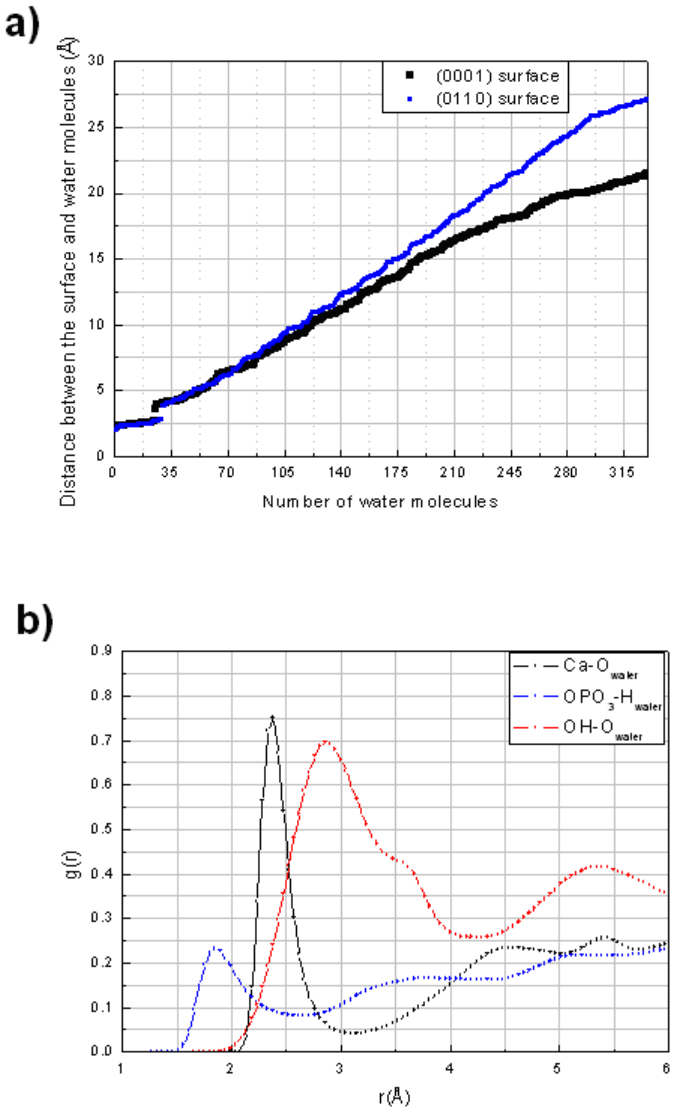


Figure 4.7 (a) Diagram showing the Ca-O_{water} distance for all water molecules (■ (0001) and ■ $(01\bar{1}0)$ HA surfaces) and **(b)** Radial distribution function between Ca-O_{water}, OPO_3-H_{water} and $^{-}$OH-O_{water} in (0001) surface.

Both hydroxyapatite surfaces under investigation are stabilised when the liquid water is introduced in the simulations. However, their relative stabilities remain the same upon hydration, with calculated surface energies of 0.37 and 0.90 Jm^{-2} for the (0001) and $(01\bar{1}0)$ surfaces respectively (Table 4.6). The adsorption of water increases the coordination of the surface ions, which has a stabilising effect

on the surfaces. The interactions of water with the surfaces were found to be from most to least important: (i) formation of strong Ca···O_{water} bonds; (ii) coordination by hydrogen atoms to surface oxygen atoms; (iii) intermolecular hydrogen-bonded networks between water molecules; and (iv) smoothing of the surface by adsorption in surface vacancies. In addition, the presence of water also induces the onset of dissolution of the surface through the escape of hydroxy groups from the materials into solution.

The formation of hydration layers on the HA surface is in agreement with previous theoretical works. For example, Zahn *et al*., have studied the interfaces between HA and water molecules by means of classical molecular dynamics simulation (Zahn & Hochrein, 2003). In their calculation, the (0001) and $(000\bar{1})$ surfaces of monoclinic hydroxyapatite were both hydrated by a very large number of water molecules, and several strongly ordered hydration layers resulted from the simulation, the first of which was characterised by a network of strong electrostatic/hydrogen bonds that reduce the water mobility. Also, Sato *et al.* (Sato et al., 2002) observed the crystalline–amorphous interfaces and the grain boundaries of sintered HA by high-resolution electron microscopy (HRTEM). When the interfaces or boundaries were parallel to $(10\bar{1}0)$, both of them showed the same atomic arrangement for the termination of the HA crystal. The crystal structure of HA was terminated at a plane crossing the hydroxyl columns on which CaII sites and PO_4 tetrahedra were located. This atomic arrangement for the $(10\bar{1}0)$ surfaces of HA probably has a very low surface energy and may be common in HA crystals formed under various conditions.

4.5 Chapter conclusions

We have used a combination of compatible potential models for the description of the interactions of the various atoms in the systems, namely the TIP3P model (Walser *et al.*, 2001) for water, and the de Leeuw (de Leeuw, 2004b) potential for hydroxyapatite. The de Leeuw potential model reproduces the structural and physical properties of the material, in agreement with experiment and the DFT calculations of the bulk material. The simulations of the dry (0001) and $(01\bar{1}0)$ surfaces have identified the (0001) surface to be the most stable, due to the combination of relative high coordination numbers of the surface species and large interlayer spacing. When hydrated, the $(01\bar{1}0)$ surface, which is highly reactive due to its large number of under-coordinated surface species, is also stabilised, but the (0001) surface remains by far the more stable surface, which agrees with experiment where the (0001) surface is expressed in the crystal morphology. The stability of the hydrated surfaces derives from direct interaction of $Ca\cdots O_{water}$ bonding, $O_{surface}\cdots H_{water}$ interactions, and $O_{water}\cdots HO$ hydrogen-bonding.

In the next chapters we will present our calculations of the adsorption of more complex molecules at hydroxyapatite surfaces, *i.e.* amino acids and peptides, which is relevant for the interaction between hydroxyapatite and biological molecules and matrices.

Chapter 5 Density Functional Theory Study of the Binding of Glycine, Proline and Hydroxyproline to Hydroxyapatite Surfaces

The investigation of the interaction of the amino acids glycine (GLY), proline (PRO), and hydroxyproline (HYP) of the collagen matrix with hydroxyapatite surfaces is our initial step to study the nucleation of hydroxyapatite at a collagen matrix. We have employed electronic structure techniques to study different modes of adsorption of the individual amino acids at the hydroxyapatite surface; in particular, those functional groups of the collagen helix that are accessible to the hydroxyapatite mineral platelets.

5.1 Amino acid molecules

Amino acids are compounds of major importance in living organisms, where they move freely in blood circulation. Each amino acid has a central carbon, called the α-carbon. This α-carbon is attached to four groups: a basic amino group (—NH_2), an acidic carboxyl group (—COOH), a hydrogen atom, and a group called a side chain (—R), which is distinctive to each amino acid (Figure 5.1 (a)). Amino acids have both amine and carboxylic acid functional groups and are therefore both an acid and a base: the amino group could win a proton and the carboxyl group could lose its hydrogen, shown in Figure 5.1(b). The net charge of the amino acid is still zero, but now it is a dipolar ion or zwitterion. The ionic state of the acidic and basic groups of the amino acids depends on the pH. Amino acids can exist as

zwitterions in solids and in polar solutions such as water, but not in the gas phase (Remko & Rode, 2006).

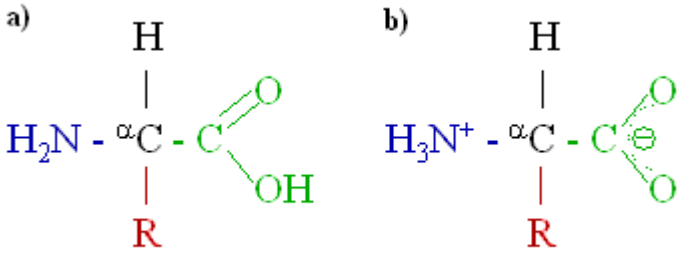


Figure 5.1 Representations of the amino acid. **a)** an amino acid has an acidic carboxyl group (—COOH), a basic amino group (—NH_2), a hydrogen atom, and a side chain (—R), all bonded to a central carbon (α-carbon). **b)** an amino acid in zwitterionic form. ($^{\alpha}C$ = grey, —COOH = green, —NH_2= blue, H = black).

The side chains vary considerably from ionized carbonyl groups with net negative charges, strongly basic group that are protonated and have a net positive charge, to uncharged polar groups, and nonpolar groups. Amino acids are usually classified by the properties of their side chain, according to their interactions with water and their charges. Hydrophobic side chains, which are nonpolar, are partially soluble in water, whereas hydrophilic side chains (acidic, basic and polar) are soluble in water.

The amino acids are the building blocks from which all proteins are made, so that the properties of the amino acid side chains determine the conformation and function of the proteins, e.g. collagen molecule. The collagen molecule is a mechanically strong, fibrous, connective-tissue protein, which is rich in glycine,

proline and hydroxyproline. Since the properties of the amino acid are so important, we have examined them more closely.

Glycine has the simplest structure of any amino acids. The small size of its side chain (one hydrogen atom) gives glycine a unique function in the structures of many proteins since it will fit in niches that can accommodate no other amino acids. Proline is the only amino acid whose side group links to the amino group and, thus, is technically an *imino* (—NH—) rather than an amino ($—NH_2$) acid. Although proline is not chemically reactive, its carbon ring restricts the geometry of the backbone chain of the protein that contains it. Hydroxyproline is incorporated into the polypeptide chain as proline and is hydroxylated after the protein has been synthesized. Hydroxyproline plays a key role in collagen stability.

The three amino acids were optimised as neutral molecules and as zwitterions, which would be their preferred configuration in water. We found that in isolation the neutral forms of the molecules were preferred over the charge-separated configuration, which rearranged upon geometry optimisation. The optimum structures of the free amino acids are largely due to the formation of intra-molecular hydrogen-bonds, which are shown in Figure 5.2. The instability of the zwitterionic structure within gas phase calculations has been reported previously for glycine in detailed *ab initio* investigations (Ding & Kroghjespersen, 1992, Yu *et al.*, 1992), and it is therefore not surprising that the same is found for proline and hydroxyproline.

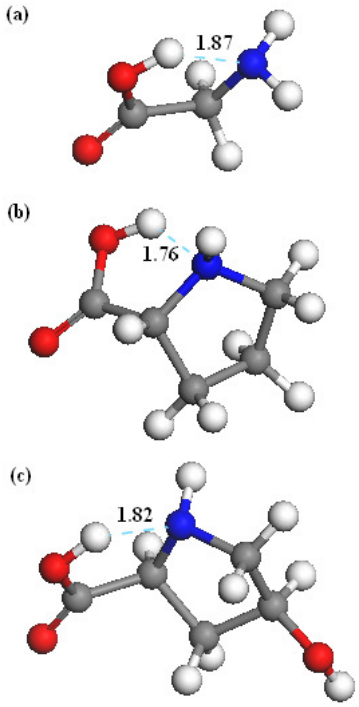


Figure 5.2 Calculated lowest energy structures of **(a)** glycine, **(b)** proline, **(c)** hydroxyproline, showing intramolecular hydrogen-bonding (C = grey, O = red, N = blue, H = white).

Although adsorption in water would probably take place with the amino acids in their zwitterionic configuration, within the protein the amino acids would not be charged either and previous simulations of the adsorption of organic molecules with amine and carboxylic acid groups to apatite and other mineral surfaces have shown that the use of neutral molecules provides quantitative insight into adsorption modes and strengths, especially in a comparison between different surfaces and adsorbates (de Leeuw & Cooper, 2004, Cooper & de Leeuw, 2004, Mkhonto *et al.*, 2006). Thus, the processes including interactions with water molecules were not considered in these calculations, although chapter 7 will build on this study to use classical methods to improve our understanding of the effect of water on the adsorption of organic molecules at the hydroxyapatite interface.

5.2 Adsorption of amino acids at HA surfaces

Pure hydroxyapatite is a stable calcium phosphate phase and the model compound for the inorganic component of bone and teeth. Clearly studies of the fundamental interactions of the three amino acids, glycine, proline, and hydroxyproline with two major hydroxyapatite surfaces can be important for the development of biomaterials for biocompatible medical implants.

In our investigation of the sorption of the amino acids at the HA surfaces, we have adsorbed one molecule per surface simulation cell at a variety of different surface sites to identify the energetically most favourable location. Experimental studies report that the primary association between amino acids and the hydroxyapatite surface is via interactions between the carboxylic acid groups and the surface calcium ions. Furthermore, amine groups could interact with oxygens of phosphate groups (Koutsopoulos & Dalas, 2000b, a, 2001). This information was useful to identify likely adsorption sites in both (0001) and $(01\bar{1}0)$ surfaces, as starting positions for our simulations. In addition, the amino acids were adsorbed at the relaxed surfaces in their optimised geometries in a number of different initial configurations, after which the complete surface/adsorbate systems were geometry optimised again and the adsorption energies calculated.

In the calculations, all atoms were allowed to translate, rotate and distort freely during the relaxation to optimise their interaction. However, for the amino acid molecule to distort significantly from its ideal structure, clearly the strength of

binding to the surface will have to outweigh the loss of energy due to the breaking of the intra-molecular hydrogen-bonded network in the free molecule.

In each case, the surface simulation cells were supercells of sufficient size (262-354 $Å^2$), containing 176 atoms and a vacuum gap of at least 50 Å, to ensure the absence of any computational artefacts due to the periodic boundary conditions parallel to the surface or across the vacuum gap. If the cells had been too small the adsorbed amino acids could have interfered with their images in the periodically repeated surface cell, which would affect the geometries and energies. The large real space supercells reduce the number of *k*-points required and only the Γ point is included for all calculations. We have first calculated the adsorption energy for the glycine/surface systems for different slab thicknesses 2×2×1, 2×2×2 and 2×2×3, but increasing the slab thickness did not affect the adsorption energies significantly, as shown in Table 5.1.

Table 5.1 Calculated adsorption energies for the glycine at different HA slabs

Adsorption energies (kJ mol^{-1})				
Surface	Termination	2x2x1 (176 Atoms)	2x2x2 (264 Atoms)	2x2x3 (528 Atoms)
(0001)	PO_4- Ca- PO_4- PO_4-Ca-Ca	-291.00	-290.82	-290.29
$(01\bar{1}0)$	OH-OH-Ca-Ca	-496.56	-499.12	-495.67

5.2.1 Glycine

The lowest energy geometry of the glycine molecule adsorbed at the (0001) surface is shown in Figure 5.3(a) (GLY-1), where we see that the glycine molecule adsorbs into the gap in the surface, which is an under-coordinated and hence reactive site. The amine group is located above a surface phosphate group, where in the bulk material a calcium ion would be sited, and the amine group's hydrogen atoms form a number of hydrogen-bonded interactions to the oxygen atoms of the phosphate group with ($NH\cdots OPO_3$) distances ranging from 1.47 Å to 1.54 Å. In addition, the oxygen atoms of the carboxylic acid group form a number of interactions with the topmost calcium ions ($Ca\cdots O$ = 2.29-2.33 Å), and as a result, the hydrogen atom of the carboxyl group migrates to the amine group. As such, the amino acid adsorbs as a zwitterion in the energetically preferred mode. The electrostatic interactions between the carboxylic acid group and calcium ions and the hydrogen bonds between amine group and surface oxygen atoms of the phosphate group give rise to strong binding of the glycine to the (0001) surface of hydroxyapatite.

The main difference between this lowest-energy structure and GLY-2 (shown in 5.3(b)) is the proton transfer from the carboxylic acid group to the amine group of glycine, rather than to a surface phosphate group, which occurs in GLY-2. GLY-3 and GLY-4 have less interaction between the carboxylic acid group and the surface calcium ions, as shown in 5.3 (c) and (d) respectively. In summary, the most favourable mode of adsorption of glycine at the hydroxyapatite (0001)

surface is in its zwitterionic form, which is in agreement with the recent computational study of Rimola *et al.*(Rimola *et al.*, 2008).

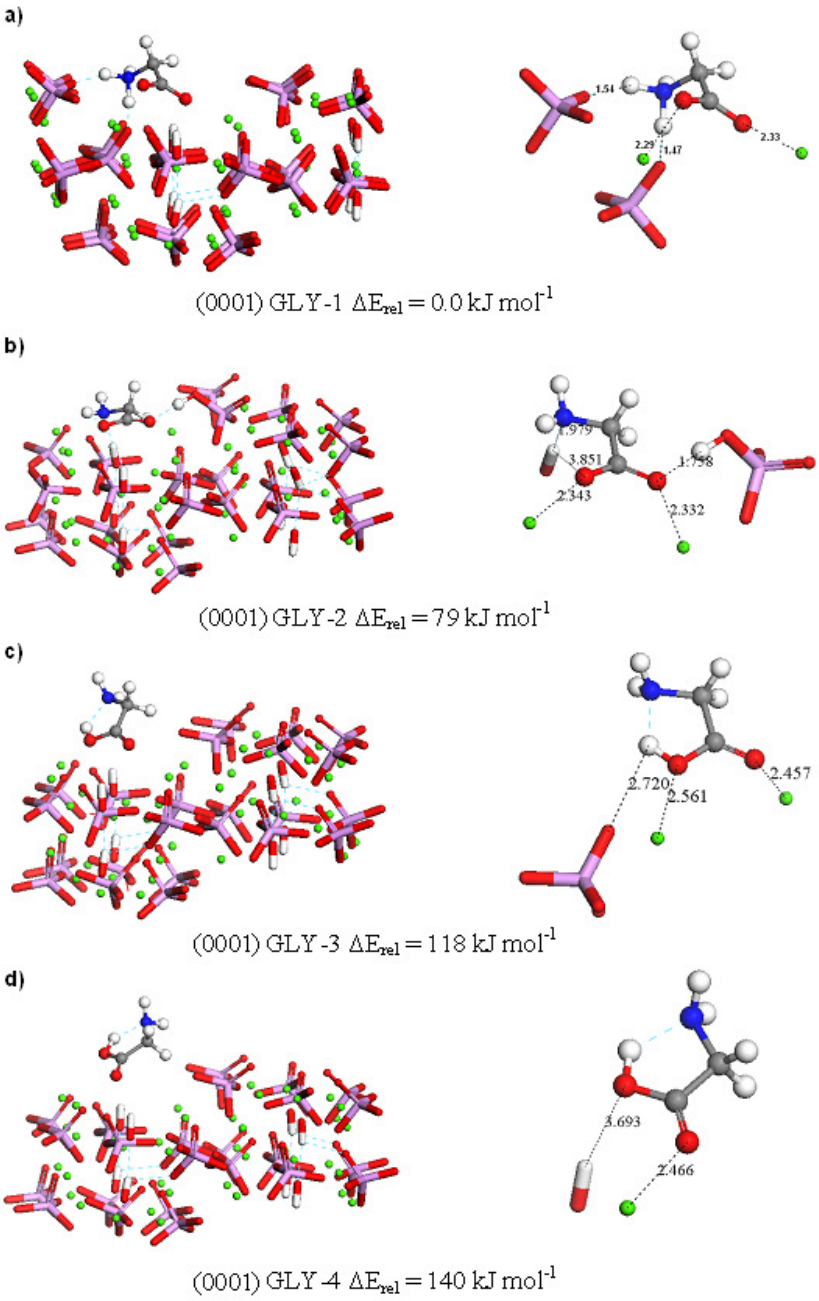


Figure 5.3 Geometry optimised structures of different configurations of glycine adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

On the $(01\bar{1}0)$ surface, the most favourable mode of adsorption for glycine is by bridging between three calcium atoms through coordination of the carboxylic acid group and the nitrogen, shown in Figure 5.4 (a) (GLY-1). Upon adsorption to the surface, the hydrogen atom of the carboxylic acid group migrates to the oxygen atom of a surface phosphate group, leaving the glycine molecule with a carboxylate group, which interacts strongly with surface calcium ions as follows: one carboxylate oxygen is coordinated to one calcium at a distance of 2.43 Å and to an other calcium at a distance of 2.39 Å, whereas the other carboxylate oxygen is coordinated to a calcium at a distance of 2.37 Å. The nitrogen atoms interact with calcium at a distance of 2.55 Å. In addition, a range of hydrogen-bonds between the hydrogen atoms of the amine group and oxygen atoms of the surface hydroxy and phosphate groups further enhance the binding between the glycine molecule and $(01\bar{1}0)$ surface ($NH\cdots OPO_3$ = 2.07-2.76 Å) in the preferred GLY-1 configuration. The GLY-2 structure (Figure 5.4 (b)) is very similar to GLY-1 (difference in energy is only 7 kJ mol^{-1}), the main difference being the more numerous interactions between the carboxylic acid group and surface calcium ions and the hydrogen-bonding between the carboxylate group and the hydroxy group present in the GLY-1 configuration.

However, in GLY-3 the proton migrates to a surface hydroxy group but this process is energetically less favourable by approximately 111.00 kJ mol^{-1}. This proton transfer leads to the formation of a molecule of water that competes with the anionic form of glycine for adsorption at the surface of hydroxyapatite. The GLY-4 is similar to GLY-3, although the positions of the amine group and

carboxylic acid group are different. Glycine is the only adsorbate in a study where this proton transfers to a surface hydroxy group (see Figures 5.4 (c) and (d)). With the other adsorbates, the carboxyl proton migrates to the oxygen atom of a surface phosphate group.

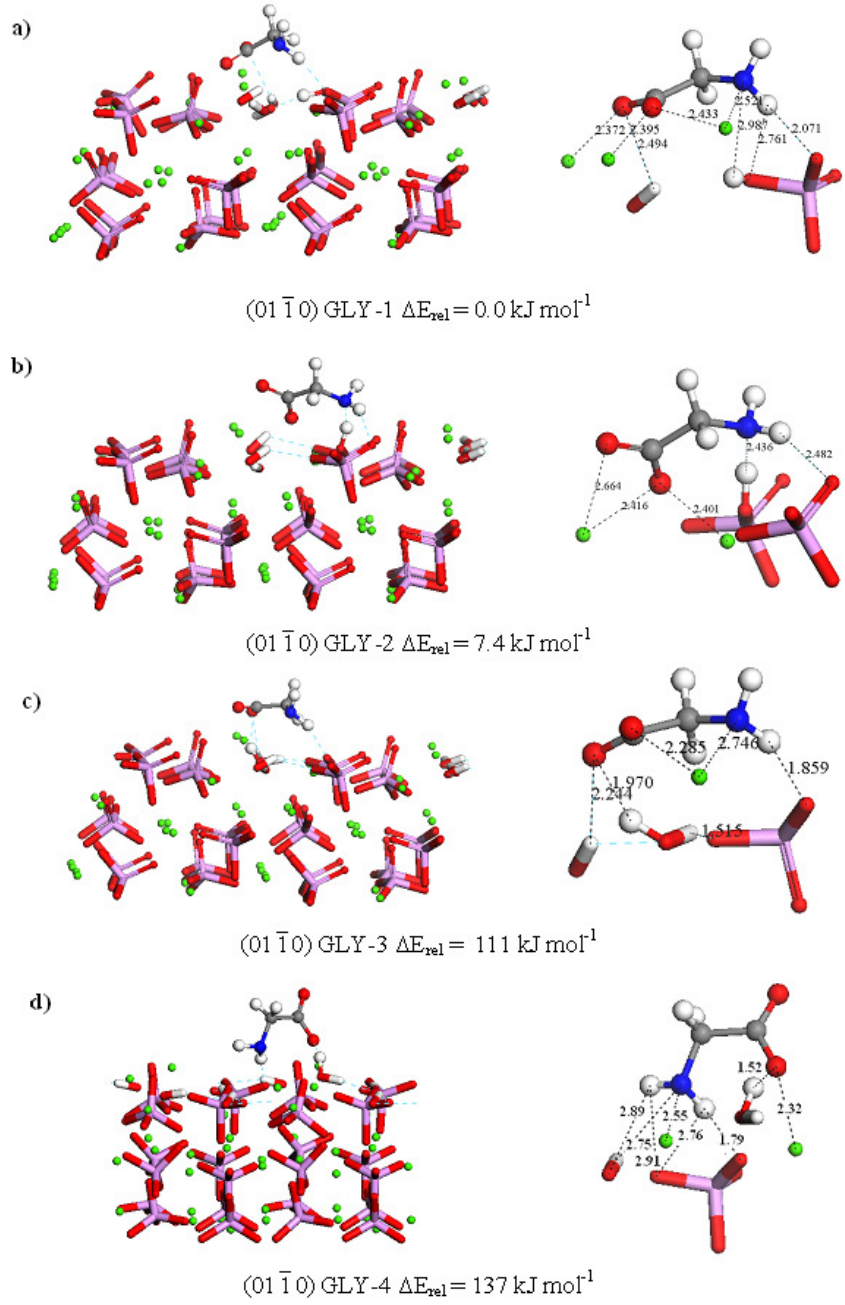


Figure 5.4 Geometry optimised structures of different configurations of glycine adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

5.2.2 Proline

Figure 5.5 (a) shows the lowest energy configuration of the proline molecule, (PRO-1) adsorbed at the (0001) surface. The hydrogen atoms of the amino group interact with the surface phosphate oxygen atoms ($NH\cdots OPO_3$ = 1.48-1.78 Å), whereas the carboxylic acid group interacts directly with two calcium ions and a hydroxy group ($COO\cdots Ca$ = 2.41-2.54 and $COO\cdots HO$ = 2.61 Å). These strong interactions between the surface and proline molecule again cause the carboxyl proton to migrate to the amine group, where the three hydrogen atoms interact with the oxygens of a surface phosphate group. The basic character of the hydroxyapatite (0001) surface is revealed by the PRO-2 configuration, where proton transfer occurs from the carboxylic acid group of proline to the basic phosphate group of the surface, but at an energy penalty of 190 kJ mol^{-1}. In this case, both its oxygen atoms of the carboxylic acid group interact with the same surface calcium atom, with Ca-O distances of 3.58 Å and 3.75 Å, they are far away from the surface than the oxygen atoms of the carboxylic acid group of the lowest energy configuration (Figure 5.5 (b)).

PRO-3 exhibits features similar to those of PRO-1, as shown in Figure 5.5(c), the main difference being the position of the amino group, which has been swapped round, and PRO-1 has more interaction between the carboxylic acid group and surface calcium ions. The hydrogen atoms of the ring in PRO-4 interact directly with the phosphate oxygen atoms at the (0001) surface ($CH\cdots OPO_3$ = 2.22-2.92 Å), while the amino group and the carboxylic acid group do not face the surface, i.e. in this configuration the proline does interact through its carboxylic acid

group. However, the energies released upon the adsorption of proline in this mode is of the order of physisorption energies (<100 kJ mol^{-1}) and thus we can not expect this configuration to be significant.

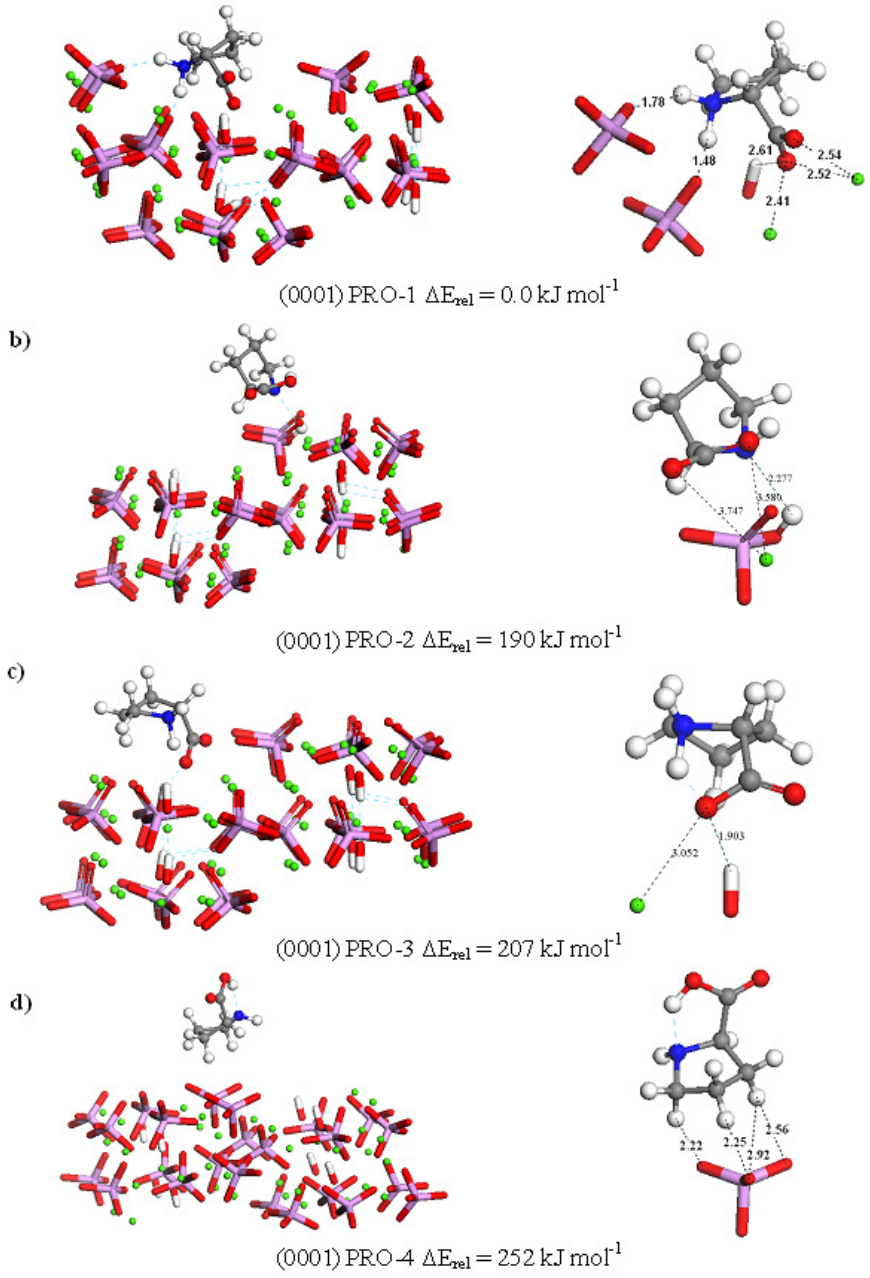


Figure 5.5 Geometry optimised structures of different configurations of proline adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

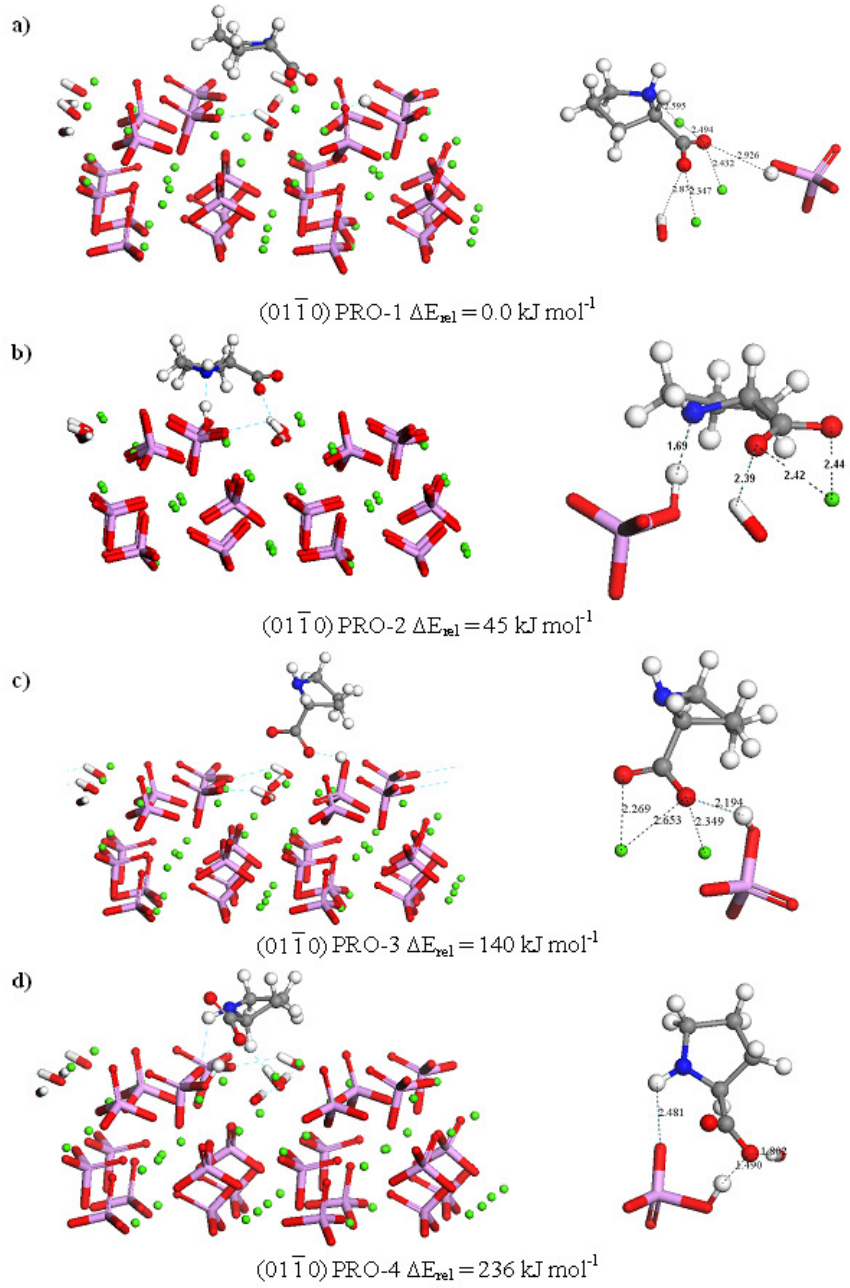


Figure 5.6 Geometry optimised structures of different configurations of proline adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

On the $(01\bar{1}0)$ surface, however, the most favourable adsorption mode for the proline molecule is through coordination of both oxygen atoms of the carboxylic acid group to surface calcium atoms, at distances of 2.35, 2.43 and 2.49 Å, and a

surface hydroxy group (CO···HO = 2.88 Å) (Figure 5.6 (a)), together with electrostatic interactions between nitrogen atoms to a surface calcium (-NH···Ca = 2.59 Å). The carboxyl proton migrates, this time first to the amine group followed by a hop to the oxygen atom of a surface phosphate group.

The PRO-2 structure shows rather similar behaviour as far as formation of one strong interaction between the carboxylic acid group and one calcium ion is concerned (Figure 5.6 (b)), except PRO-2 has fewer interactions with surface calcium ions. Both the PRO-3 and PRO-4 structures (shown in Figure 5.6 (c) and (d)) interact less with the surface, mainly coordinating by the carboxylic acid group to surface calcium ions. In summary, proline is adsorbed at the $(01\bar{1}0)$ surface in its anionic form due to proton transfer toward the surface, confirming the higher reactivity of this surface compared to the (0001) surface.

5.2.3 Hydroxyproline

Hydroxyproline adsorbs in a similar way to proline, which is not surprising as the structures of the two molecules are similar, where only one ring hydrogen of proline is replaced by a hydroxy group to form hydroxyproline. The substitution adds to the number and strength of interactions between the adsorbate and the surface species. For example, in hydroxyproline both carboxylic acid and hydroxy groups can interact with surface calcium ions, whereas their hydrogens can form numerous hydrogen-bonded interactions to phosphate and hydroxy oxygen atoms of the surface.

Figure 5.7 (a) (HYP-1) shows the most favourable mode of adsorption of hydroxyproline to the (0001) surface, where the interaction with the surface occurs via the hydrogen and oxygen atoms of the hydroxy group and hydrogens of the amine group ($OH\cdots OPO_3$ = 1.59-2.71 Å, $OH\cdots HO$ = 2.14 Å and $NH\cdots OPO_3$ = 1.46-2.96 Å). The carboxylic acid group, again, interacts with surface calcium ions ($COO\cdots Ca$ = 2.29-2.79 Å), as well as oxygen atoms of the hydroxy group ($HO\cdots Ca$ = 2.91 Å) and the carboxyl proton migrates to the amine group.

In the HYP-2 and HYP-3 structures, the proton of the carboxylic acid group migrates to the oxygen of the surface phosphate group, where the main difference between the two is that the hydroxyproline adsorbs into the gap in the surface in the HYP-2 configuration. In these cases, hydroxyproline acts as an acid tightly bound to the basic sites (PO_4^{-3}) of the hydroxyapatite surface.

HYP-4 exhibits features different to rest of the configurations. Figure 5.7 (d), shows the adsorption in the HYP-4 configuration, where the interaction with the surface occurs via the hydrogen atoms of the hydroxy group and $>CH_2$ ($OH\cdots OPO_3$ = 1.64 Å, $CH\cdots OPO_3$ = 2.69-2.99 Å), without participation of the amino or carboxylic acid groups**.**

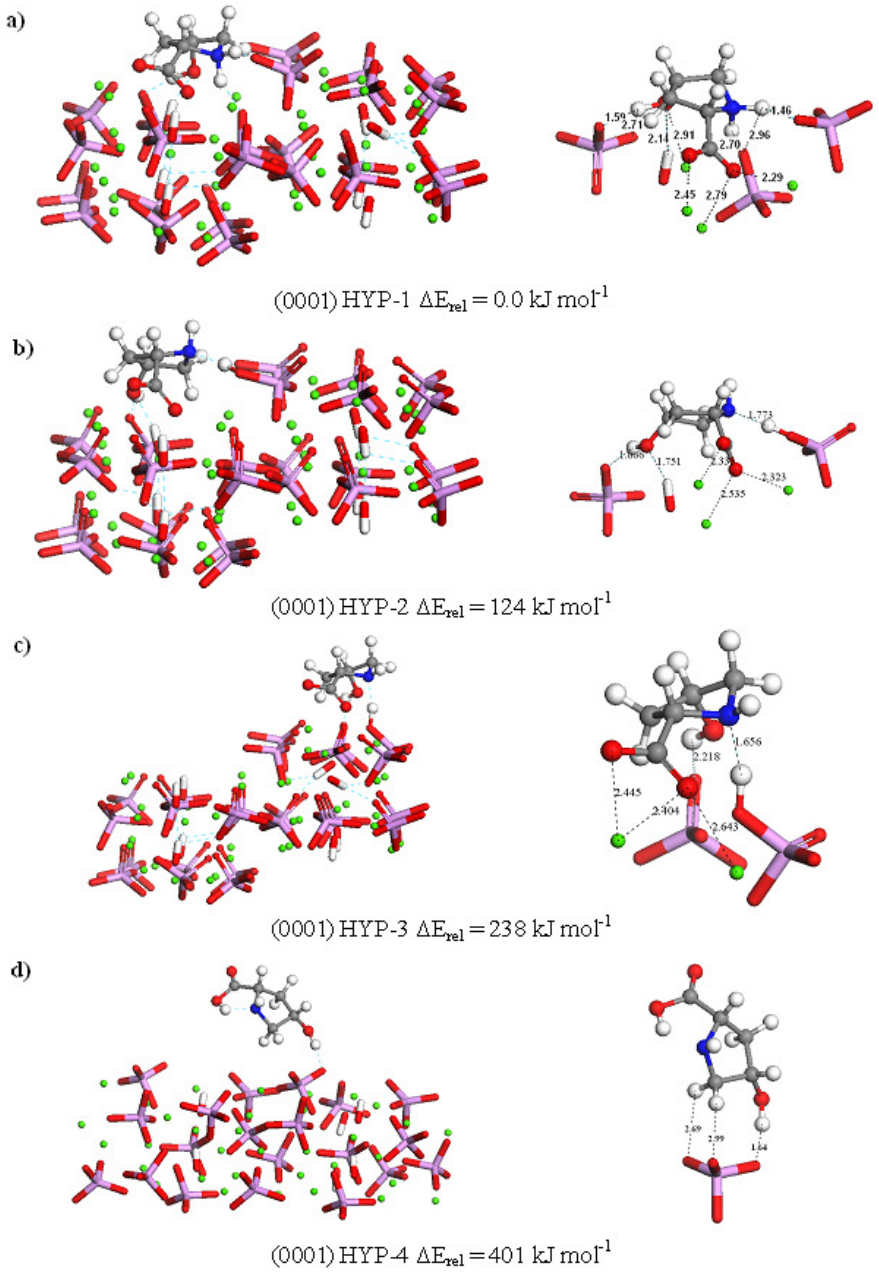


Figure 5.7 Geometry optimised structures of different configurations of hydroxyproline adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

On the $(01\bar{1}0)$ surface, hydroxyproline adsorbs almost flat onto the surface (HYP-1), where a number of close interactions are formed between the molecule's nitrogen atom and oxygen atoms of the carboxylic acid group and three surface

calcium ions (COO···Ca = 2.34-2.38 Å, N···Ca = 2.61 Å). In addition, hydrogen-bonding takes place between oxygen atoms of a surface phosphate group and the hydrogen atoms of the molecule's −OH and $>CH_2$ groups (-OH···OPO_3 = 1.78-2.99 Å, CH···OPO_3 = 2.04-2.19 Å). As in proline, the carboxylic acid proton migrates to an oxygen atom of the surface phosphate group, and further hydrogen-bonding occurs between the oxygen atoms of the carboxylate group and the hydrogen atoms of the two topmost hydroxy groups (COO···HO = 2.83-2.85 Å).

In HYP-2, the hydroxyproline again coordinates to the $(01\bar{1}0)$ surface via its carboxylic oxygen atoms to surface calcium atoms at Ca···O distances of between 2.38, 2.42 and 2.44 Å. The hydroxyl hydrogen atoms are hydrogen-bonded to oxygen atoms of surface carbonate groups, with -OH···OPO_3 distances of between 1.61 and 2.73 Å.

In the HYP-3 configuration, hydroxyproline adsorbs almost flat onto the surface, with the main difference to the HYP-1 structure being the number of the interactions between the carboxylic acid group and surface calcium ions. However, in HYP-4, the adsorbate molecule is almost perpendicular to the surface, forming interactions to one calcium ion through two of its carboxylic acid groups, but it does not form hydrogen-bonding interactions through its hydroxy and amino groups.

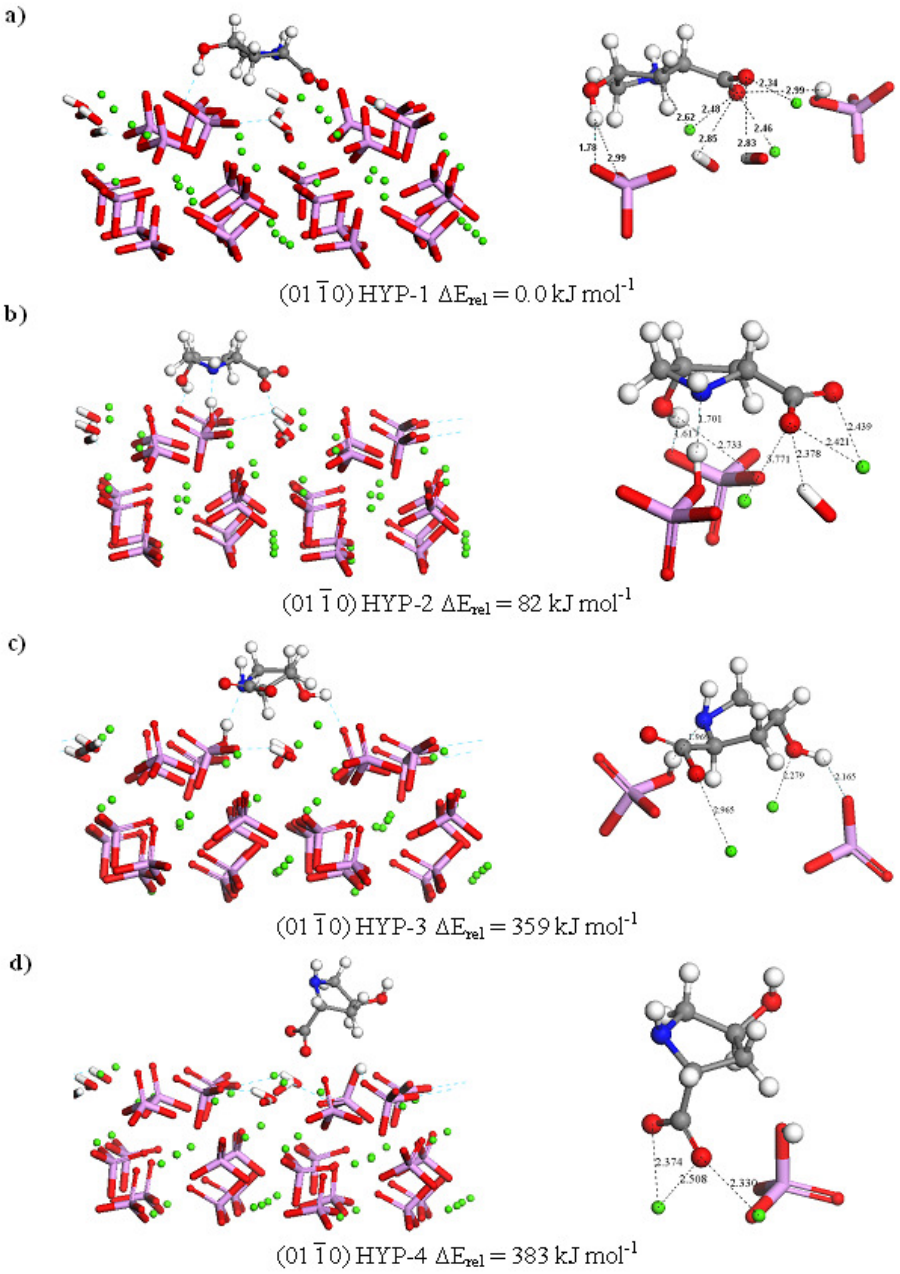


Figure 5.8 Geometry optimised structures of different configurations of hydroxyproline adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances and corresponding relative energies (ΔE_{rel}) (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

5.3 Adsorption Energies

The adsorption energies for all amino acids at the surfaces are listed in Table 5.2 Experimental studies (Koutsopoulos & Dalas, 2000b, a) have shown that amino acids interact strongly with HA, which is borne out by our calculations. However, all adsorbates bind much more strongly to the $(01\bar{1}0)$ than to the (0001) surface. The (0001) surface is a stable surface with only a few dangling bonds before adsorption, and as a result the adsorbates have little effect on the surface structure. However, the many interactions between adsorbate and surface lead to the large energies released upon adsorption. In summary, all amino acids are adsorbed at the (0001) surface in their zwitterionic form, whereas the anionic form can be adsorbed but at an energy penalty.

As discussed, the $(01\bar{1}0)$ surface is less stable than the (0001) surface, which is due to a larger number of under-coordinated surface species with dangling bonds, making the $(01\bar{1}0)$ surface more reactive than the (0001) surface. Saturation of the dangling bonds through the adsorption of the amino acids, as well as the proton transfer occurring upon adsorption from the adsorbate to the surface, leads to the large chemisorption energies released upon adsorption at the $(01\bar{1}0)$ surface. The process of adsorption at the $(01\bar{1}0)$ surface always includes chemical reactions in the way of proton transfer to the surface, which creates charge separation between the surface and adsorbate, leading to strong ionic bonding between the carboxylate group and surface calcium ions, explaining the high adsorption energies for this surface.

Table 5.2 Calculated adsorption energies of amino acids at HA surfaces

Structure		Adsorption energies (kJ mol^{-1})	
		(0001)	$(01\bar{1}0)$
GLY	1	-291.00	-496.56
	2	-211.81	-489.14
	3	-173.18	-385.56
	4	-150.76	-359.92
PRO	1	-322.30	-554.50
	2	-132.29	-509.77
	3	-115.32	-414.04
	4	-70.58	-318.23
HYP	1	-507.65	-609.79
	2	-383.02	-527.76
	3	-270.00	-250.62
	4	-106.73	-226.62

When we compare the adsorption energies for the different amino acids (Table 5.2), we see that, in general, the energies released by adsorption of proline and

hydroxyproline are larger than glycine, even though the glycine molecule is more flexible. As the carboxylic acid and amine functional groups are present in all those adsorbates, the differences in binding strengths and modes of adsorption between the glycine and the proline or hydroxyproline must be due to the presence of the extra carbon and hydrogen atoms (and hydroxy group), which increase the interactions to the surface. This strong interaction is important, as the rings of the proline and hydroxyproline in the collagen strands are positioned on the outside of the triple helix and hence are most accessible to calcium ions and phosphate groups nucleating at the template (Prockop & Kivirikko, 1995).

Sorption of hydroxyproline is energetically the most favourable. The enhanced interaction of this molecule with the HA surfaces compared to proline is due to the extra hydroxy functional group. The hydrogen atom of the hydroxy group is flexible and thus can easily form interactions to surface oxygen atoms. In addition the electrostatic interaction between this polar group and the surface increases the binding strength.

5.4 Chapter conclusions

In this chapter we have described the results of our investigation into the adsorption of the three amino acids glycine, proline and hydroxyproline, which are major constituents of the collagen I protein, at two important hydroxyapatite surfaces, the (0001) and $(01\bar{1}0)$. Our simulations show that the strength of interaction of the amino acid molecules with the surfaces depends both on the

stability of the surface and the capability of the amino acid molecules to form multiple interactions with the surface species, particularly if they can bridge between two or more surface calcium ions. The primary association between amino acids and the HA surfaces is via interactions between $–COO^-$ and surface-bound Ca^{2+}, but the additional interactions of the side groups lead to significant variations in their affinities, consistent with experimental results (Koutsopoulos & Dalas, 2000b, a, 2001).

Our work shows a distinct preference by the three amino acids for binding to the $(01\bar{1}0)$ surface, which is required if the bone mineral platelets are to become elongated in the c-direction and to express the $(01\bar{1}0)$ surface, rather than the (0001) surface, as is observed in the morphology of the natural bone platelets. The (0001) surface is the thermodynamically preferred surface morphology but this surface is not expressed significantly in the natural bone tissue, which grows from the collagen matrix. All three amino acids adsorb strongly to the hydroxyapatite $(01\bar{1}0)$ surface, which indicates that their functional groups should be good sites for the nucleation and growth of this surface at the collagen fibrils.

In experimental studies, Dalas and co-workers (Koutsopoulos & Dalas, 2001) have found that glycine is a very weak inhibitor, and this may be the explained by absence of a side group, which could give rise to an interaction with the hydroxyapatite surface. Also, they tested hydrophobic amino acids such as proline as possible inhibitors of hydroxyapatite crystallization, and these were found to

have a low adsorption affinity toward the hydroxyapatite surface (Koutsopoulos & Dalas, 2000b, Tanaka *et al.*, 1989). However, hydroxyproline appeared to be an inhibitor, which can be explained by the presence of the hydroxyl group in the side chain (Koutsopoulos & Dalas, 2000a). All these studies are in agreement with our results from DFT calculations.

Chapter 6 Theoretical Study of the Binding of Collagen Peptides to the Hydroxyapatite Surfaces

In the previous chapter we have investigated the electronic and structural properties of the functional groups of the three amino acids, which are accessible to the hydroxyapatite mineral when it is deposited at the collagen triple helix. In this chapter, we consider four peptides containing a hydrophobic side group (PRO), with an uncharged polar side group (GLY and HYP), and with a charged polar side group (LYS and HYL). The structural properties of the isolated amphiphilic PRO-HYP-GLY and HYP-PRO-GLY peptides, and hydrophobic PRO-LYS-GLY, and PRO-HYL-GLY peptides, as well as the interaction of these peptides with the (0001) and $(01\bar{1}0)$ surfaces of hydroxyapatite have been simulated, using density functional theory methods. These peptides have been chosen because they give the opportunity to investigate the effect of hydroxylation and of charged side group in the strength of peptide/hydroxyapatite interaction.

6.1 Introduction

A free amino acid can react with another and join in a peptide chain; the link between them is called peptide bond. An amino group from one amino acid forms an amide bond with the carboxyl of the second residue with loss of a molecule of water (Figure 6.1). The linked amino acids are called amino acid residues, and the free amino group and carboxylate group at opposite ends of the polypeptide chain

are called amino terminus (N-terminus) and carboxyl terminus (C-terminus), respectively. The size of polypeptide chains is variable, and depends on the number of the amino acid residues forming the peptide, i.e. chains that contain two amino acid residues are named dipeptides, three are called tripeptide, those that contain fewer than 50 are oligopeptides, longer chains are polypeptides, whereas the term protein is reserved for larger chains.

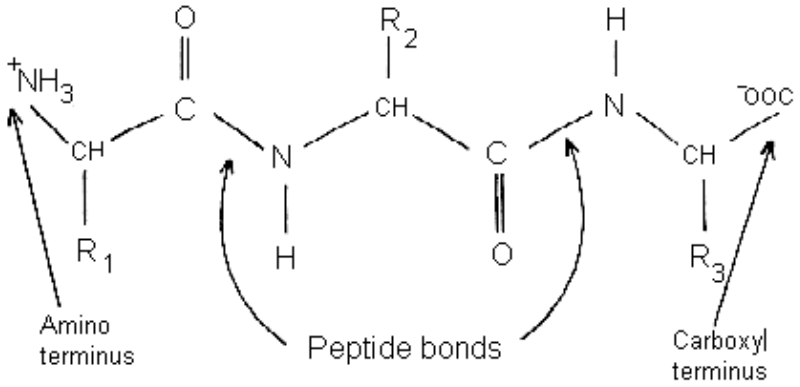


Figure 6.1 Peptide link between three amino acids. The result is a tripeptide in which the amino acid residues are linked by a peptide bond.

We have used density functional theory techniques to study the binding process of four representative tripeptides to the (0001) and $(01\bar{1}0)$ surfaces of hydroxyapatite. Since, we are particularly interested to investigate whether this small peptide might have a growth modifying effect, as well as the effect of hydroxylation and the charge in the side group on the strength of interaction with the surfaces, the peptide chains were constructed through a combination of the amino acid residues proline (PRO), hydroxyproline (HYP), glycine (GLY), lysine (LYS), and hydroxylysine (HYL), which are shown in Figure 6.2. Glycine, proline and hydroxyproline, as described in the section 5.1, are neutral amino

acids. The side chain of lysine contains an amine (NH_2) attached to the terminal side carbon. The side-chain amino group of the lysine is a strong base, and, at pH 7.0, the predominant ionic species is the ammonium ion, —NH_3^+. Lysine is covalently modified to 5-hydroxylysine in the same way as the proline. We have therefore designed a number of different tripeptides containing hydrophobic, uncharged polar side groups, and charged polar side groups, namely HYP-PRO-GLY, PRO-HYP-GLY, PRO-LYS-GLY, and PRO-HYL-GLY.

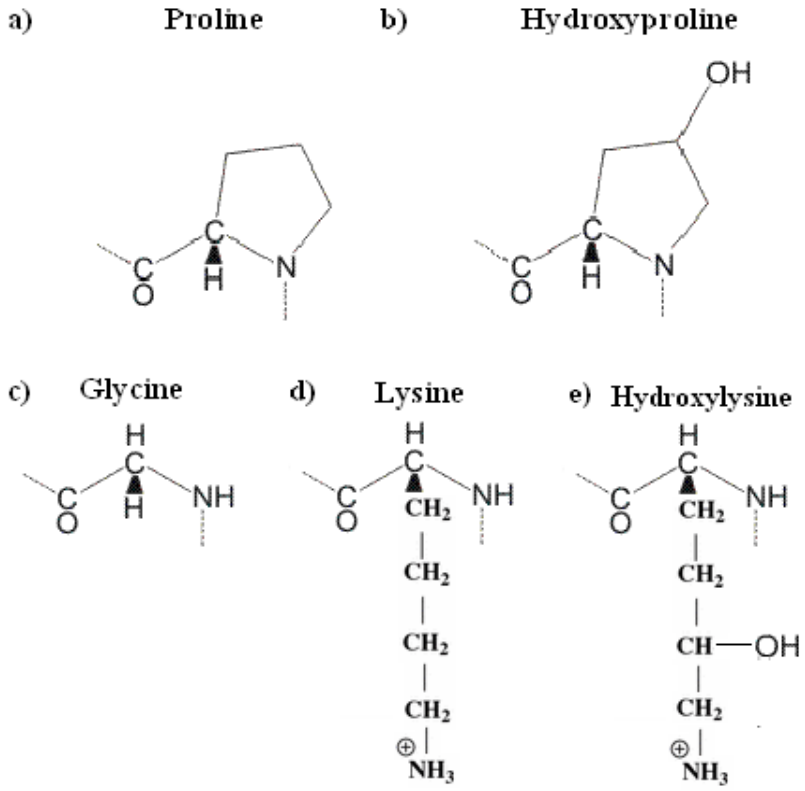


Figure 6.2 Scheme of amino acids: **a)** Proline (hydrophobic side group), **b)** Hydroxyproline (uncharged polar side group), **c)** Glycine (uncharged polar side group), **d)** Lysine (charged polar side group), and **e)** Hydroxylysine (charged polar side).

The (0001) and $(01\bar{1}0)$ surfaces of hydroxyapatite were simulated as described in section 4.1.1. The peptides were adsorbed in various configurations, used

information the chapter 4 to predict possible adsorption site. All calculations have been performed using the SIESTA program (Ordejon *et al.*, 1996).

6.2. Results and Discussion

We first modelled the peptide molecules as zwitterions, which is their preferred form in aqueous environment and, in contrast to isolated amino acids (Almora-Barrios *et al.*, 2009), it is also the most stable form in vacuo. The conformational structures of the peptides, which are shown in Figure 6.3, are largely due to the formation of intra-molecular hydrogen-bonds.

The geometries of HYP-PRO-GLY (6.3a) and PRO-HYP-GLY (6.3b) are similar, although there is a small but signicant difference between their total energy of 7.8 kJ mol^{-1}. In both peptides, the oxygen atoms in the carbonyl and carboxylate groups form hydrogen bonds with the hydrogen atoms of the amine groups. However, in PRO-LYS-GLY (6.3(c)), the formation of hydrogen bridges is closest to the amino group of the side chain, whereas in PRO-HYL-GLY (6.3(d)) hydrogen bonds are formed between the amine group and the hydroxy group of the side chain, and the hydroxy group and carboxylate oxygens, where the hydroxy thus forms a bridge between amino and carboxylate groups.

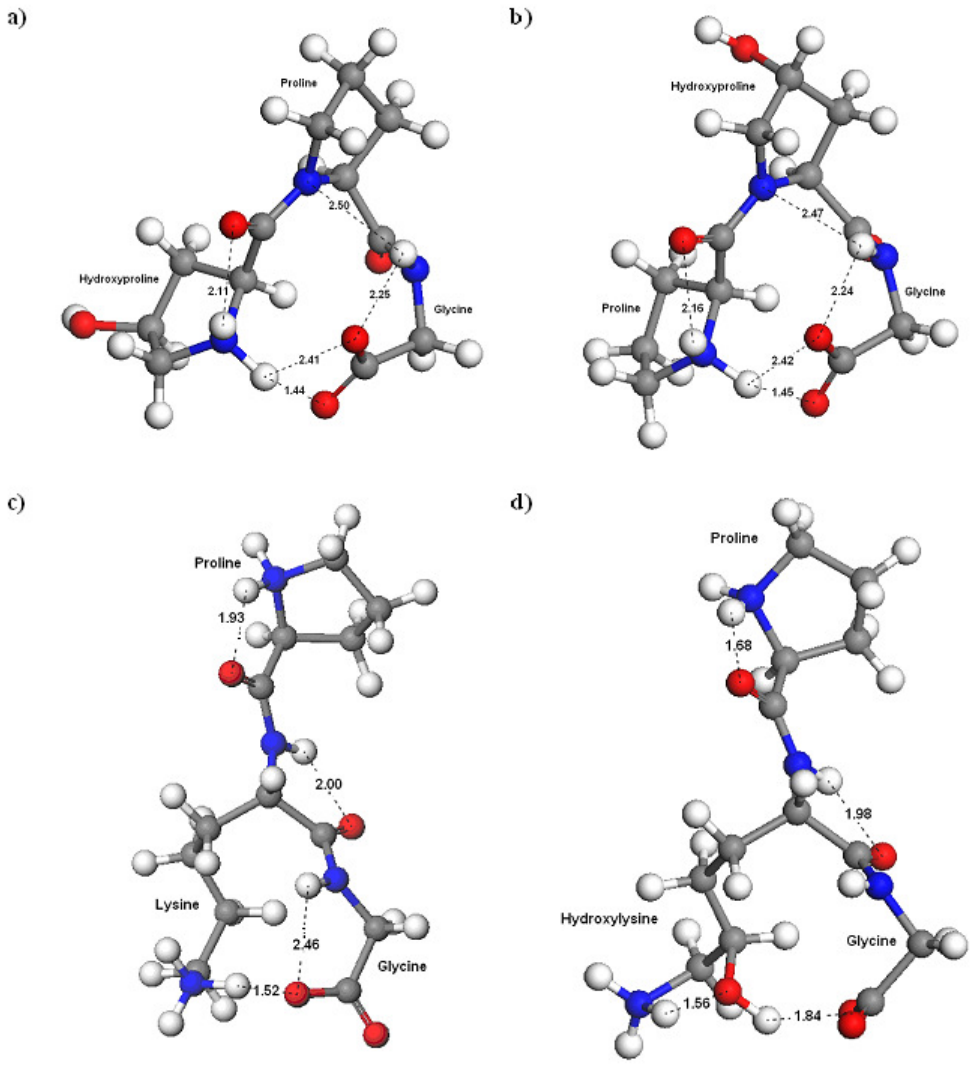


Figure 6.3 Optimised structure of **a)** [$^{+}H_2N$-HYP-PRO-GLY-COO^{-}], **b)** [$^{+}H_2N$-PRO-HYP-GLY-COO^{-}], **c)** [$^{+}H_2N$-PRO-LYS-GLY-$COO^{-}]^{+}$, and **d)** [$^{+}H_2N$-PRO-HYL-GLY-$COO^{-}]^{+}$, showing the occurrence of intramolecular hydrogen bonding (C = gray, O = red, N = blue, and H = white).

Having optimised the free structure of peptides, they were next placed above the surfaces. The calculations show that all peptides form strong interactions with the hydroxyapatite surfaces, which confirms the affinity of the —NH_3^{+} (or >NH_2^{+}) and —COO^{-} groups for hydroxyapatite, consistent with the estimates reported for different protein/HA systems (Chen *et al.*, 2008, Gibson *et al.*, 2006, Goobes *et al.*, 2007, Long *et al.*, 2001, Makrodimitris *et al.*, 2007, Shaw *et al.*, 2000,

Capriotti *et al.*, 2007). In the next section, we will discuss in details the adsorption of each peptide with the (0001) and $(01\bar{1}0)$ surfaces.

6.2.1 Adsorption of HYP-PRO-GLY peptide

The lowest energy geometry of the HYP-PRO-GLY peptide adsorbed at the (0001) surface is shown in Figure 6.4, where one oxygen atom of the carboxylate group interacts simultaneously with two calcium ions (Ca···O = 2.41-2.61 Å), and the second oxygen of the carboxylate group interacts with a further calcium ion (Ca···O = 2.63 Å). One carbonyl group interacts with other surface calcium ions (Ca···O = 2.40 Å), while a range of hydrogen-bonds between the hydrogen atoms of the amine groups and oxygen atoms of the phosphate groups further enhance the binding between the peptide and the (0001) surface (NH···OPO_3 = 1.4-2.7 Å).

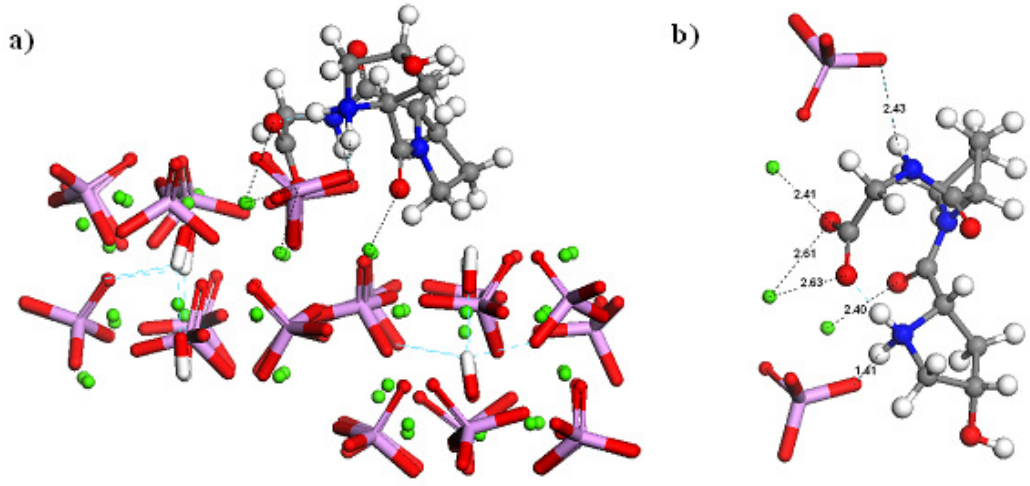


Figure 6.4 Lowest-energy structure of [$^{+}H_2N$-HYP-PRO-GLY-COO^{-}] adsorbed at the hydroxyapatite (0001) surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

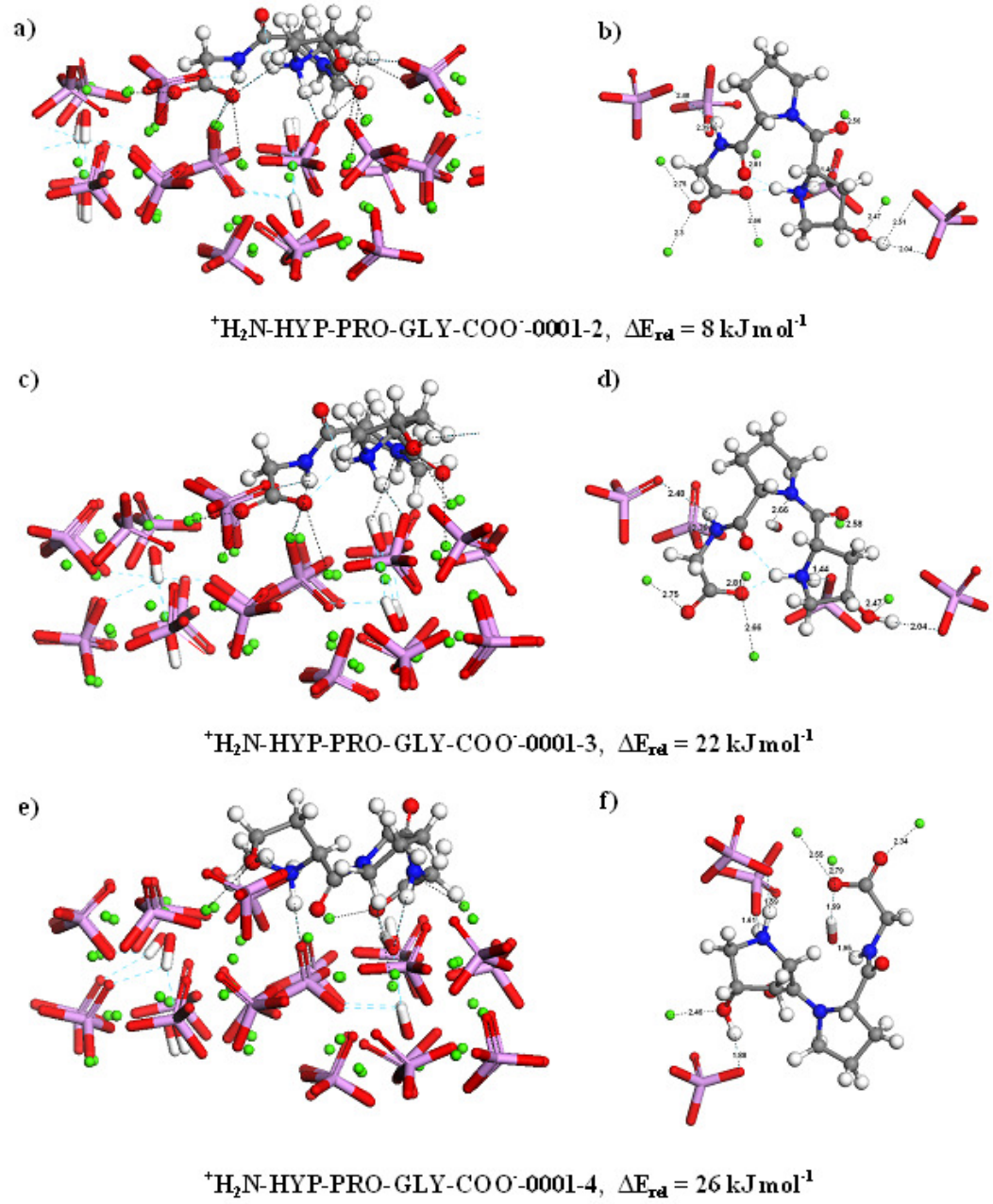


Figure 6.5 Geometry optimised structures of the different $^{+}H_2N$-HYP-PRO-GLY-COO^{-} adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

Other less favourable adsorption configurations are shown in Figure 6.5. In all cases, the configuration of the adsorbed peptide to the surface is in similar way, such as, mayor interaction is –COO^{-} group interacting with surface calcium ions

and the $>NH_2^+$ group forms hydrogen-bonding to oxygen atoms of the surface phosphate groups. Also, the HYP-PRO-GLY peptide is adsorbed at the hydroxyapatite (0001) surface in its zwitterionic form, similar to the adsorption of the amino acids (glycine, proline and hydroxyproline) with the same surface(Almora-Barrios *et al.*, 2009).

Figures 6.6 and 6.7 show the peptide molecule adsorbed at the $(01\bar{1}0)$ surface. In Figure 6.6, the oxygen atoms of the carboxylate group are shared among two surface calcium ions ($Ca\cdots OCO^-$ = 2.38-2.88 Å, $Ca\cdots O{=}C$ = 2.31, 2.76 Å), whereas the nitrogen atom interacts with calcium at a distance of 2.70 Å. In addition, the hydrogen atoms of the amine group form hydrogen-bonds with the oxygen atom of the phosphate group ($NH\cdots OPO_3$ = 2.3 Å). The mechanism of adsorption on this surface differs from the (0001) surface, as now the hydrogen atom of the amine group ($>NH_2^+$) migrates to the oxygen atom of a surface phosphate group, leaving the peptide negatively charged.

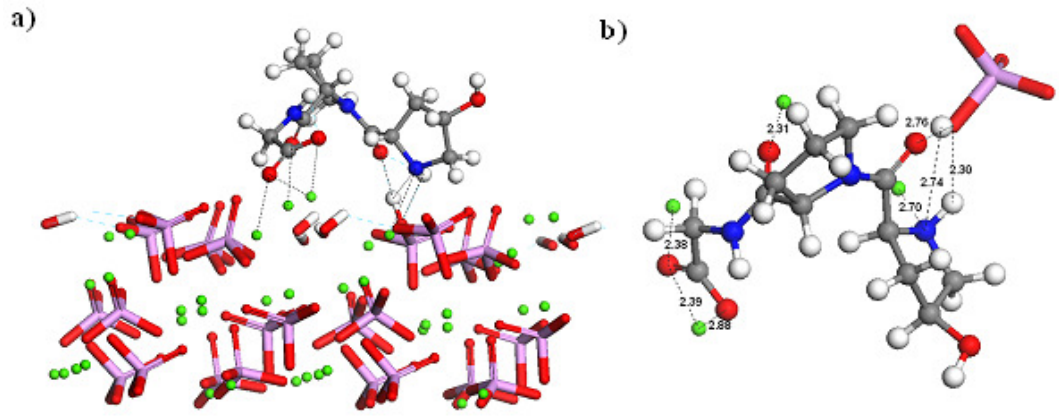


Figure 6.6 Lowest-energy structure of [^+H_2N-HYP-PRO-GLY-COO^-] adsorbed at the hydroxyapatite $(01\bar{1}0)$ surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

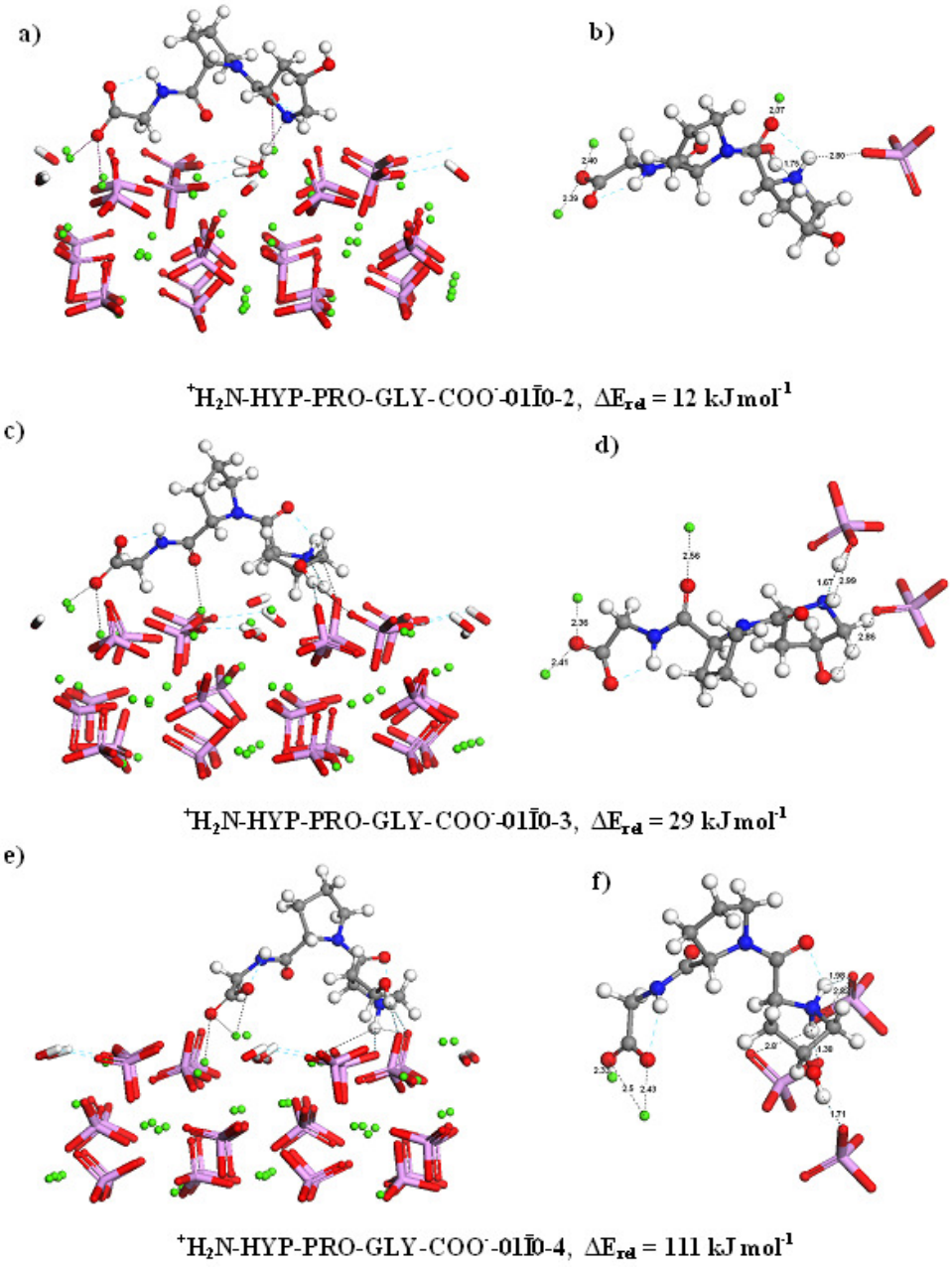


Figure 6.7 Geometry optimised structures of the different $^{+}H_2N$-HYP-PRO-GLY-COO^{-} adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

The second lowest-energy configuration is similar to the lowest-energy configuration, but has fewer interactions between the oxygen atoms of the carboxylic acid group and surface calcium ions. However, there is one case where

the peptide is adsorbed as a neutral zwitterion (shown in Figure 6.6), which is ~36 kJ mol^{-1} higher in energy than the anionic form. In the least favourable configuration, the proton migrates to a surface hydroxy group, similar to what was found for glycine adsorption(Almora-Barrios *et al.*, 2009), to form a molecule of water that competes with the anionic form of the peptide for adsorption on the surface of hydroxyapatite, but at an energy penalty of 111 kJ mol^{-1} compared to the lowest-energy configuration.

Since both $^{-}COO^{-}$ and $>NH_2^{+}$ groups are involved in peptide bonds in the chains of the collagen molecule, we next changed the position of amino acids residues in the peptide in order to study the effect of side chain groups in the interaction with the surface of hydroxyapatite.

6.2.2 Adsorption of PRO-HYP-GLY peptide

The most stable configuration of the PRO-HYP-GLY peptide adsorbed at the (0001) hydroxyapatite surface is shown in Figure 6.8 (a, b).

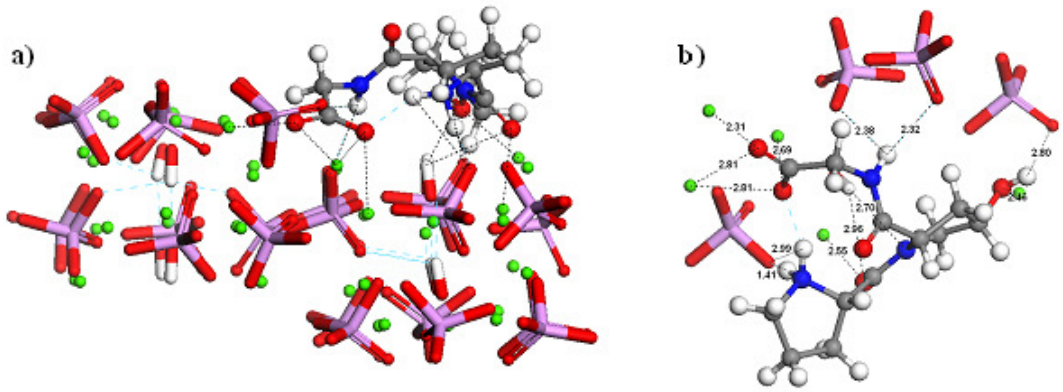


Figure 6.8 Lowest-energy structure of [$^{+}H_2N$-PRO- HYP-GLY-COO^{-}] adsorbed at the hydroxyapatite (0001) surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

Again, the peptide is located on the edge of the step on the surface, with the proline and hydroxyproline residues accommodated within the gap. The oxygen atoms of the carboxylate and carboxyl groups interact with surface calcium ions ($Ca\cdots OCO^-$ = 2.31-2.81 Å, $Ca\cdots O{=}C$ = 2.55 Å), with hydrogen-bonding taking place between hydrogen atoms of the amine group and phosphate and hydroxy oxygen atoms ($NH\cdots OPO_3$ = 1.41-2.99 Å). The central hydroxyproline residue appears to be in close contact with the surface, where a number of interactions are formed between the hydroxy group and surface species ($OH\cdots OPO_3$ = 2.8 Å, and $HO\cdots Ca$ = 2.46 Å).

In the second most favourable configuration, the hydrogen atom of the amine group migrates to the nitrogen atom of a surface phosphate group, leaving the peptide in its anionic form. However, this configuration is 40 kJ mol^{-1} higher in energy than the adsorption of the peptide as a zwitterion. Both the third and fourth configurations (see the Figure 6.9) are similar to lowest-energy configuration (Figure 6.8 (a, b)), which forms, however, more interactions between the carboxylic acid group and surface calcium ions and hydrogen bonds.

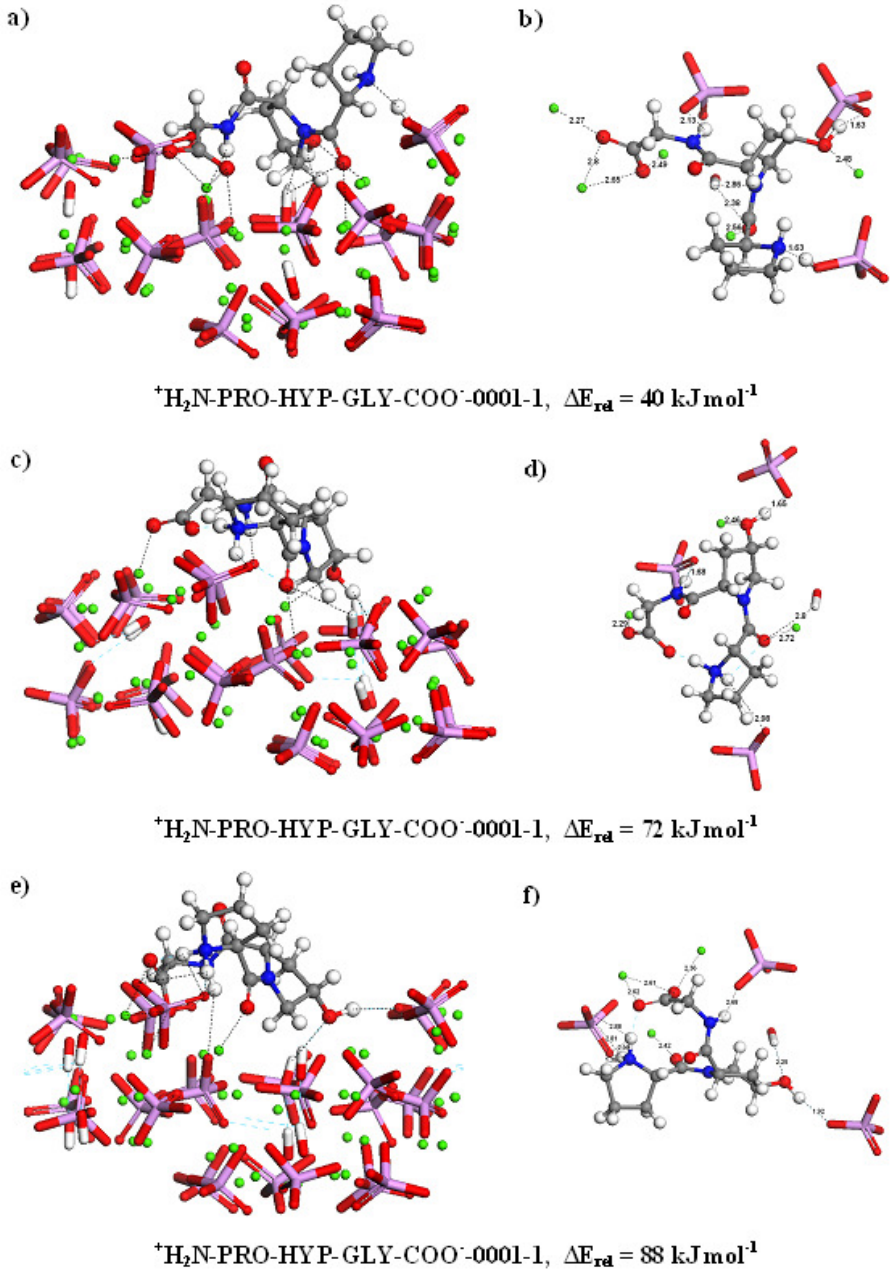


Figure 6.9 Geometry optimised structures of the different $^{+}H_2N$-PRO-HYP-GLY-COO^{-} adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

Figure 6.10 shows the most favourable configuration of the PRO-HYP-GLY peptide adsorbed at the $(01\bar{1}0)$ surface. The amine proton migrates to the oxygen atom of a surface phosphate group, and one oxygen of the carboxyl group is

shared among two calcium ions, whereas the nitrogen atom interacts with calcium at a distance of 2.51 Å. In addition, the oxygen of one carbonyl group is engaged in an interaction with the calcium ion and the other forms hydrogen bonds with the surface hydroxy group ($Ca^{\cdots}OCO^{-}$ = 2.39 Å, $Ca^{\cdots}O{=}C$ = 2.53 Å, $-C{=}O\cdots HO^{-}$ = 2.62 Å, $NH\cdots OPO_3$ = 2.77 Å, $NH\cdots OH^{-}$ = 2.85 Å).

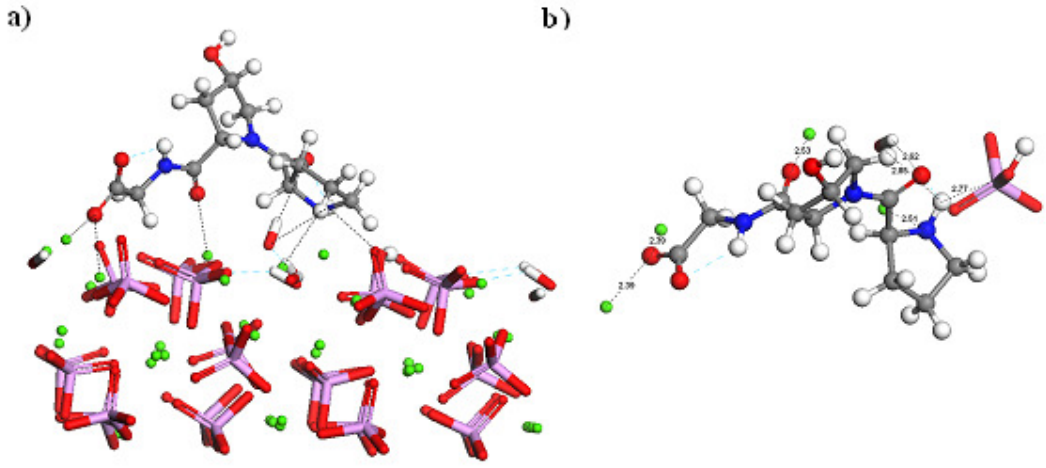


Figure 6.10 Lowest-energy structure of [$^{+}H_2N$-PRO- HYP-GLY-COO^{-}] adsorbed at the hydroxyapatite $(01\bar{1}0)$ surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

In the second and third most stable configurations (Figure 6.11), the peptide is adsorbed in its zwitterionic form but at an energetic cost of 26 and 89 kJ mol^{-1} respectively. However, in the least favourable configuration, the proton migrates to a surface hydroxy group, with the formation of a molecule of water that competes with the peptide for adsorption on the surface of hydroxyapatite, leading to an energetically even less favourable adsorption mode.

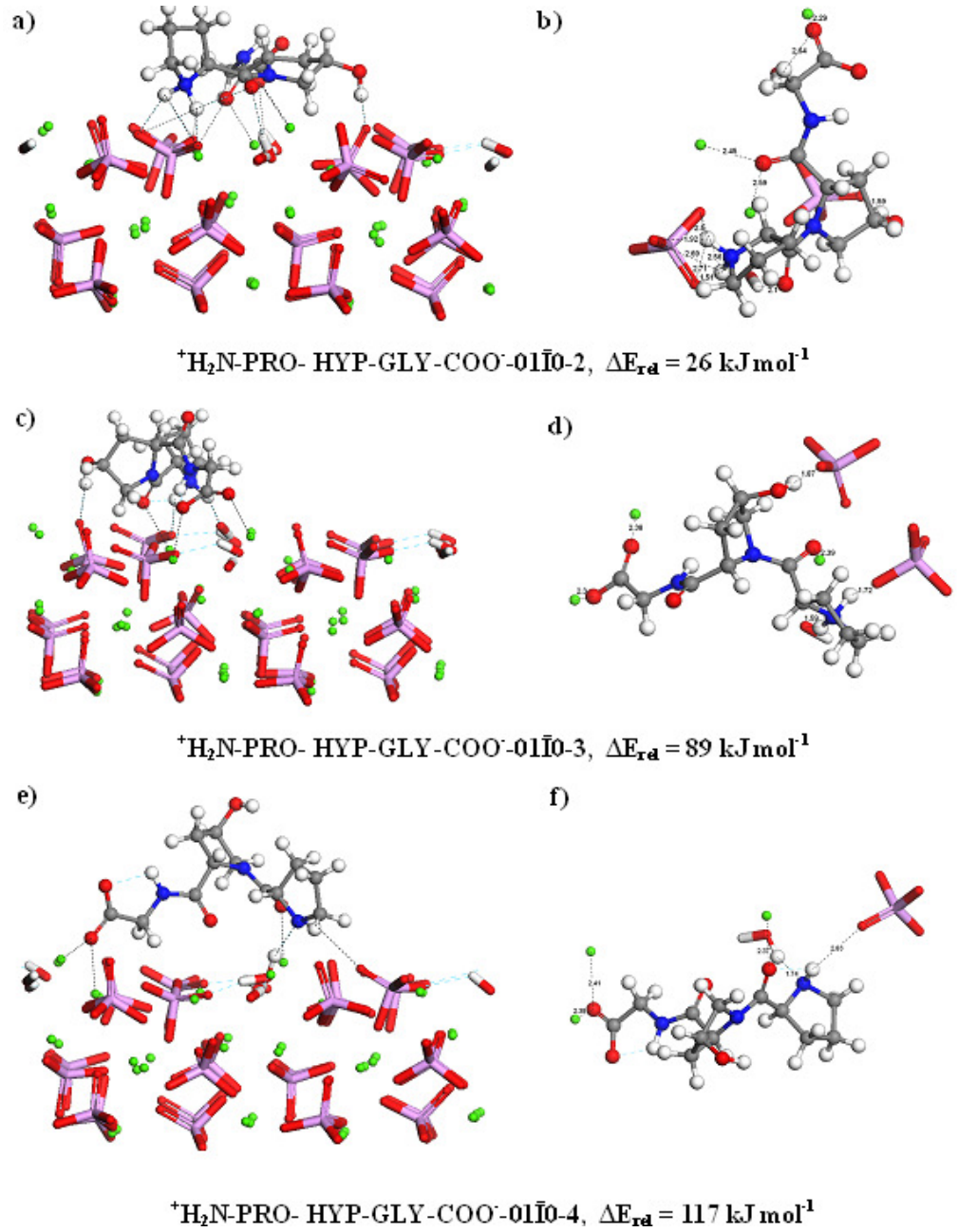


Figure 6.11 Geometry optimised structures of the different $^{+}H_2N$-PRO-HYP-GLY-COO^{-} adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

The primary association between the PRO-HYP-GLY peptide and the hydroxyapatite surfaces is again via $^{-}COO^{-}$ and $>NH_2^{+}$ groups, suggesting that amino side chains carrying either acidic or basic residues are likely to interact

strongly with hydroxyapatite surfaces. We next studied the peptide containing lysine in the PRO-LYS-GLY sequence, which has a positively charged side group, thus affording a direct comparison of the surface sorption affects of hydroxyproline and lysine.

6.2.3 Adsorption of PRO-LYS-GLY peptide

Again, the most favourable mode of adsorption for the PRO-LYS-GLY peptide with the 0001) surface (see Figure 6.12) is through a number of interactions between the oxygen atoms of the peptide's carboxylate and carbonyl groups and surface calcium ions (Ca···OCO^- = 2.46-2.65 Å, and Ca···O=C = 2.49, 2.80 Å). In addition, hydrogen-bonding takes place between hydrogen atoms and the oxygen atoms of the surface phosphate group (NH···OPO_3 = 1.52-2.94 Å), whereas the positive charge associated with the $–NH_3^+$ of the side group of the lysine residue has increased the adsorption of the peptide to the surface. As a result the hydrogen atom of the amine group of the lysine migrates to the oxygen atom of the surface phosphate group, leaving the peptide charge-neutral.

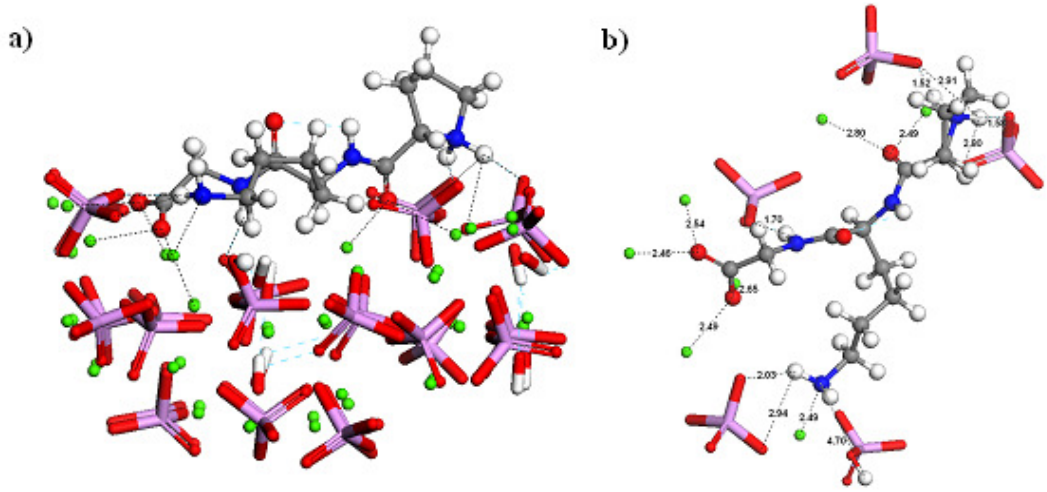


Figure 6.12. Geometry optimised structures of the different [$^{+}H_2N$-PRO-LYS-GLY-COO^{-}]$^{+}$ adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

The second configuration is 58 kJ mol^{-1} less favourable compared to the lowest energy configuration, where the main difference is that now the peptide is adsorbed in its original cationic form. The third configuration also leaves the peptide in its cationic form, but in a different orientation. The fourth configuration exhibits features similar to the lowest-energy geometry, but the peptide has lost the interactions between the proline residue and the surface species (all shown in Figure 6.13).

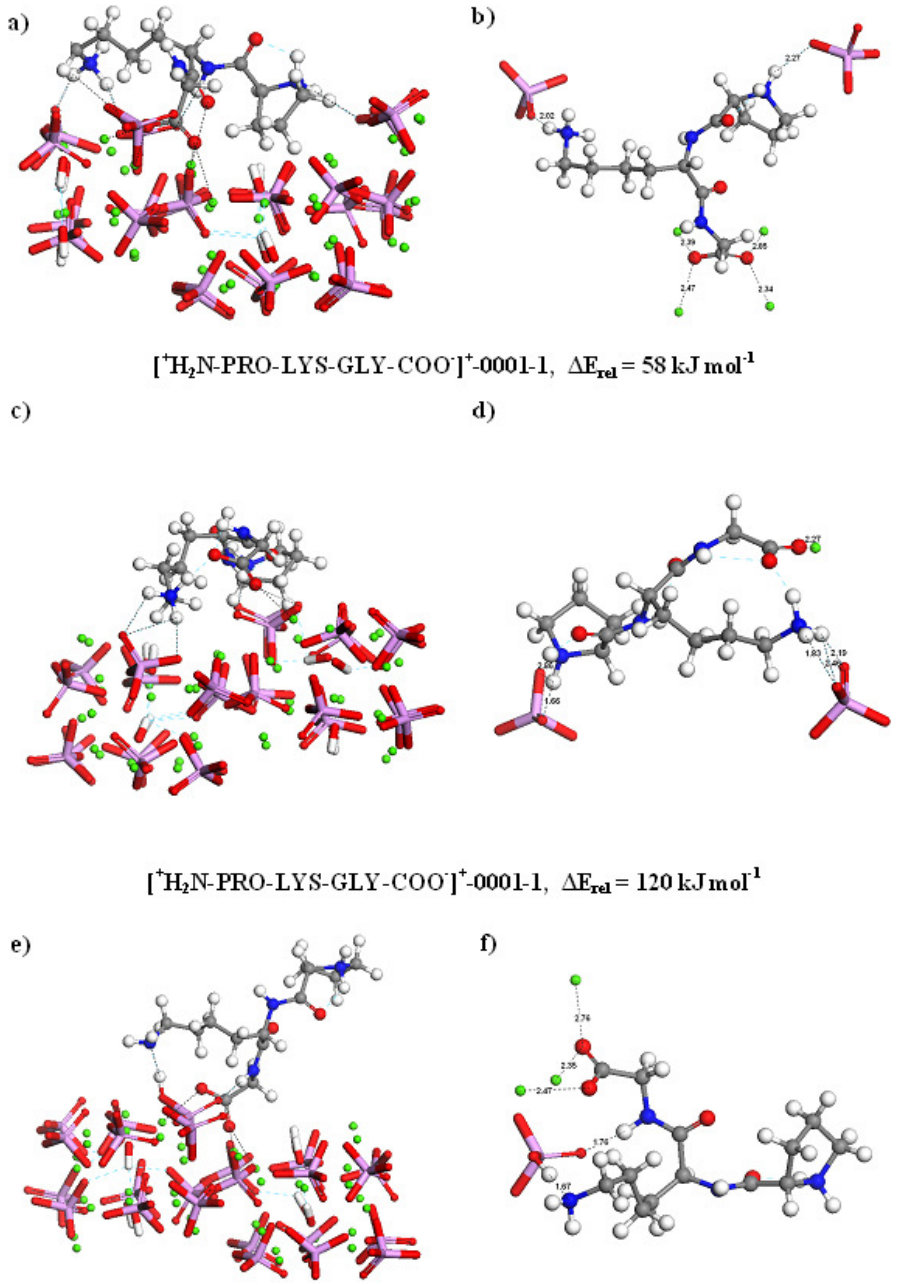


Figure 6.13. Geometry optimised structures of the different $[^{+}H_2N\text{-PRO-LYS-GLY-}COO^{-}]^{+}$ adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

On the $(01\bar{1}0)$ surface, the most favourable mode of adsorption for the peptide is though its carboxylate and amine groups, shows in Figure 6.14. Upon adsorption to the surface, the hydrogen atoms of the amine groups of both lysine and proline

migrate to the oxygen atoms of the surface phosphate groups, and the peptide is adsorbed at the surface in its anionic form. One carboxylate oxygen is coordinated to calcium at a distance of 2.38 Å and to another calcium at a distance of 2.52 Å, whereas the other carboxylate oxygen is coordinated to a calcium at a distance of 2.34 Å. Also, the carbonyl oxygen and amine nitrogen form a number of interactions with the topmost calcium ions (Ca···O=C = 2.45, 2.62 Å, and Ca···N = 2.69 Å), as well as a range of hydrogen-bonded interactions through the hydrogen atoms of the amine group (NH···OPO_3 = 1.76-2.12 Å).

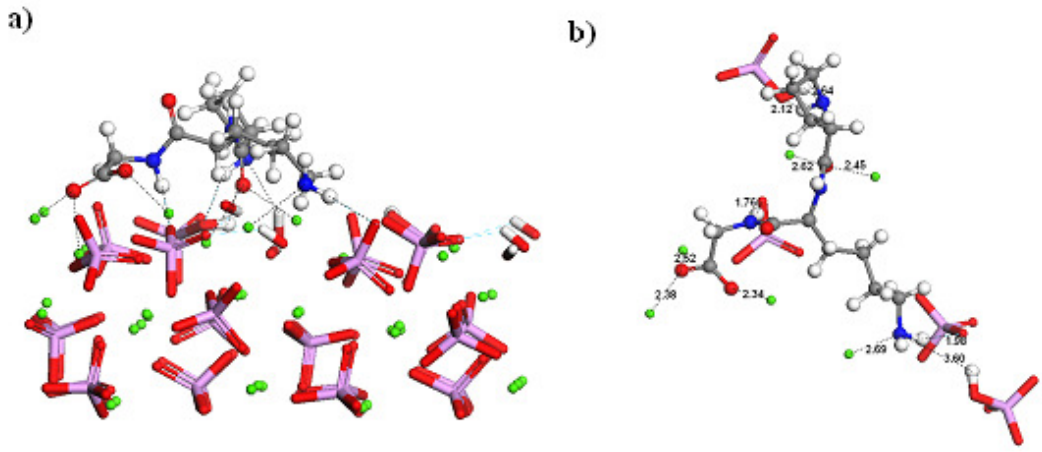


Figure 6.14 Lowest-energy structure of $[^{+}H_2N\text{-PRO-LYS-GLY-}COO^{-}]^{+}$ adsorbed at the hydroxyapatite $(01\bar{1}0)$ surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

The second configuration (shown in Figure 6.15 (a, b)) is very similar to the first (but higher in energy by 37 kJ mol^{-1}), with the main difference being the more numerous interactions between carbonyl oxygen, amine nitrogen and calcium ions and hydrogen bonding in the most stable geometry. The third configuration has lost the interaction between the amine group of the proline residue and the surface, as shown in the Figure 6.14 (c, d). In the last configuration, we see that

the proline residue is located away from the surface, so that the peptide interacts only through the carboxylate group of the glycine and amine group of the lysine. However, in all geometries the amine proton of the lysine residue migrates to an oxygen atom of surface phosphate group.

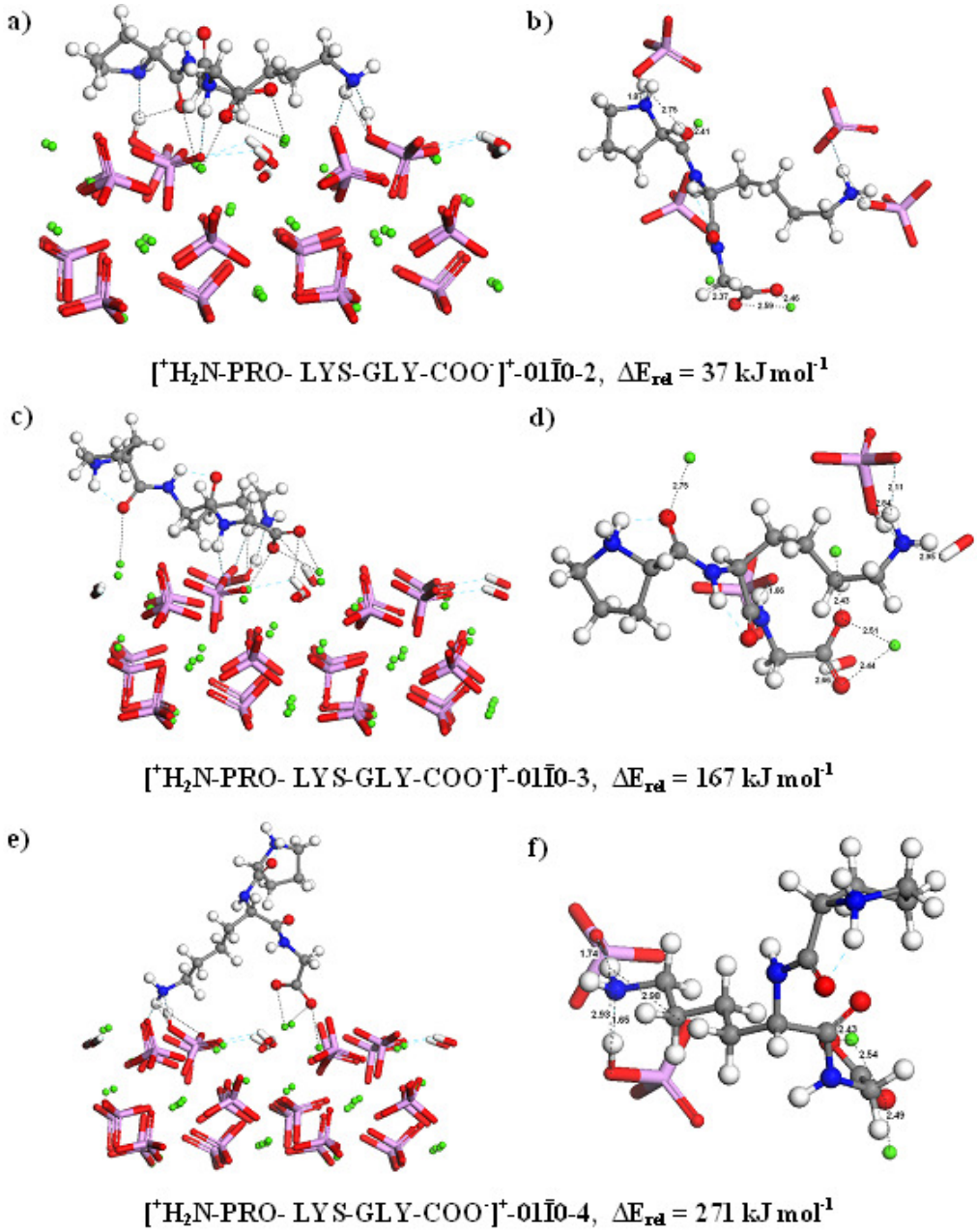


Figure 6.15 Geometry optimised structures of the different $[^{+}H_2N\text{-PRO-LYS-GLY-COO}^{-}]^{+}$ adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

6.2.4 Adsorption of PRO-HYL-GLY peptide

The hydroxylation of the lysine residue adds to the number and strength of the interactions between the peptide and the surface species. These interactions involve the formation of hydrogen bonds between the hydroxy group of the hydroxylysine residue and phosphate oxygen atoms, whereas its hydroxy oxygen can interact with surface calcium ions.

Figure 6.16 shows the lowest energy configuration on the (0001) surface, where the interaction with the surface occurs mainly via oxygen atoms of carboxylate and carbonyl groups, and hydrogens of the amine groups (Ca···OCO^- = 2.41-2.61 Å, Ca···O=C = 2.61 Å, NH···OPO_3 = 1.69, 2.15 Å). The oxygen atoms of the HLY hydroxy group interact with surface calcium ions (HO···Ca = 2.40 Å) whereas the hydrogen atom of the hydroxy group forms hydrogen-bonding interactions to a phosphate oxygen atom (-OH···OPO_3 = 1.59 Å).

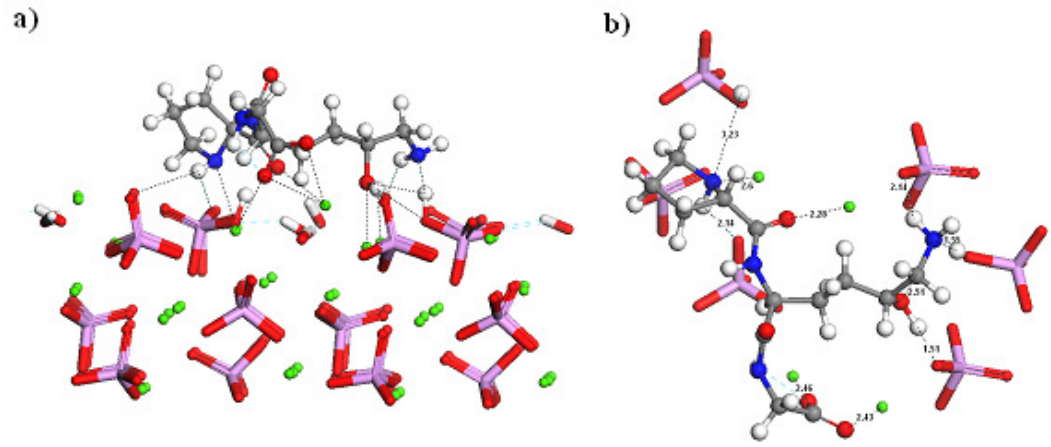


Figure 6.16 Geometry optimised structures of the different [^+H_2N-PRO-HYL-GLY-COO^-]$^+$ adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances and adsorption energies (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

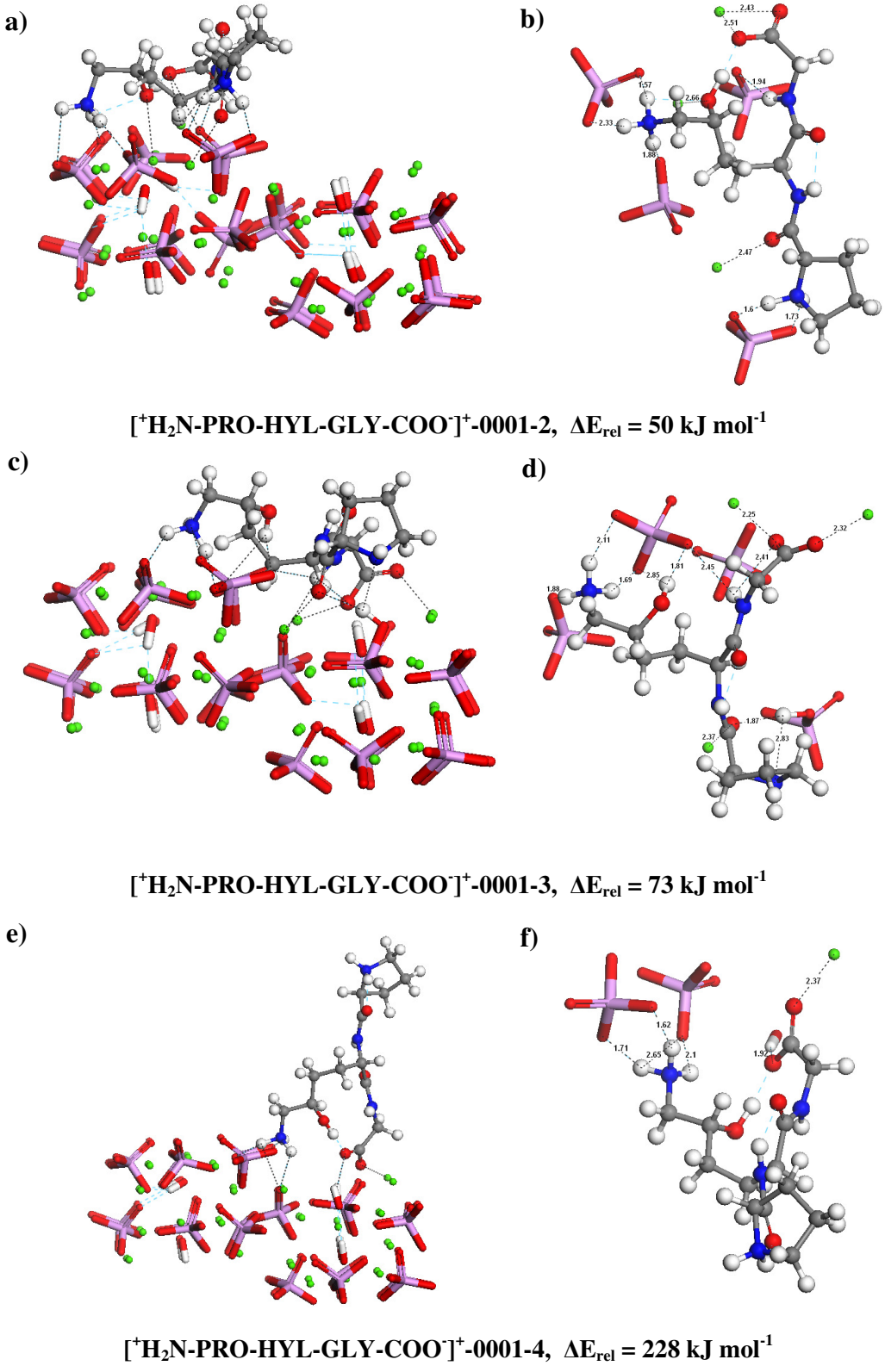


Figure 6.17. Geometry optimised structures of the different [$^{+}H_2N$-PRO-HYL-GLY-COO^{-}]$^{+}$ adsorbed at the hydroxyapatite (0001) surfaces, showing interatomic distances and adsorption energies (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

In all configurations, electrostatic attractions are likely to be important between the surface calcium ions and oxygen atoms of the carboxylate and carbonyl groups, as well as between the oxygen atoms of the surface phosphate groups and the positively charged $-NH_3^+$ of the side group of the hydroxylysine residue (see Figure 6.17). However, only in the lowest energy configuration does the amine proton of the hydroxylysine residue migrate to the oxygen atom of a surface phosphate group.

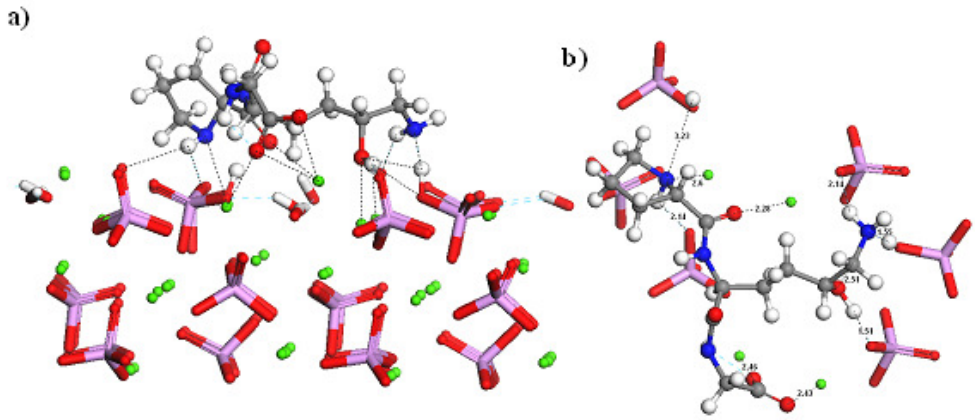


Figure 6.18 Lowest-energy structure of $[^+H_2N\text{-PRO-HYL-GLY-}COO^-]^+$ adsorbed at the hydroxyapatite $(01\bar{1}0)$ surface: **a)** overall and **b)** detailed view, showing interaction distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

The lowest-energy geometry of the PRO-HYL-GLY peptide adsorbed at the $(01\bar{1}0)$ surface is shown in Figure 6.18. The peptide adsorbs in a bidentate fashion, coordinating via both oxygen atoms of the carboxylate group to the same surface calcium atom, with Ca···O distances of 2.43 Å and 2.53 Å, whereas the oxygen atoms of its carbonyl and hydroxy groups interact at distance of 2.28 Å and 2.60 Å respectively. The hydroxy and amine hydrogen atoms are both weakly hydrogen-bonded to surface oxygen atom, with (-OH···OPO_3) distances of 1.51

and ($NH\cdots OPO_3$) 2.14-2.57 Å. As in the PRO-LYS-GLY, the hydrogen atoms of the amine group of both hydroxylysine and proline migrate to oxygen atoms of surface phosphate groups, leaving the peptide negatively charged.

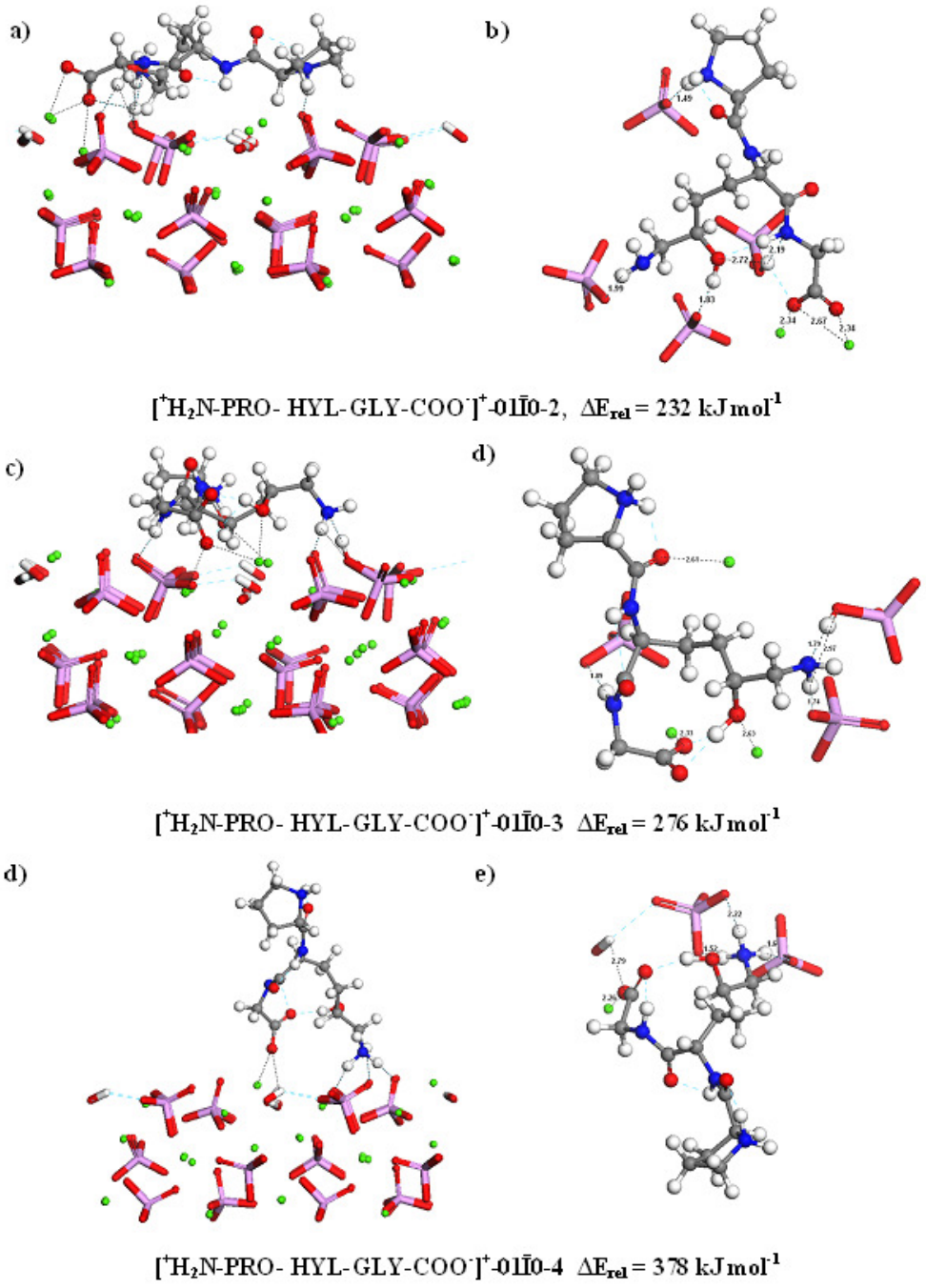


Figure 6.19 Geometry optimised structures of the different [^+H_2N-PRO- HYL-GLY-COO^-]$^+$adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces, showing interatomic distances (Ca = green, P = purple, C = gray, O = red, N = blue, and H = white).

In the second configuration, the peptide is adsorbed as a neutral zwitterion (shown in Figure 6.19 (a, b)), and only a hydrogen atom of the amine in the side group of the hydroxylysine residue migrates, this time first to a carboxylate group, followed by a hop to the oxygen atom of a surface phosphate group. The third configuration is similar to the second configuration, but with fewer interactions to surface species. The last configuration is the only one in this study, where proton transfer between peptide and $(01\bar{1}0)$ surface is not observed. However, the peptide interacts strongly with surface calcium ions and forms hydrogen-bonds between the hydrogen atoms in the amine group of the side of the hydroxylysine residue and oxygen atoms of the surface phosphate groups (see the Figure 6.19).

6.2.5 Adsorption Energies

The adsorption energies of the peptides to the hydroxyapatite surfaces, calculated according to equation 3.3, are presented in Table 6.1. When we compare the adsorption energies, we see that the energies change with the position of the constituent amino acid residues in the peptide, as well as with the type of side group. This behaviour is not surprising as the side groups determine the number and strength of the interactions between the peptides and the surfaces species. All peptides interact strongly with the two HA surfaces; however, this interaction is stronger on the $(01\bar{1}0)$ surface than the (0001) surface. The $(01\bar{1}0)$ surface is less stable than the (0001) surface and the adsorption process on the $(01\bar{1}0)$ surface always includes chemical reactions; both factors which lead to the high adsorption energies calculation for this surface.

Table 6.1 Calculated Adsorption Energies of Peptides at HA Surfaces

Structure		Adsorption energies (kJ mol^{-1})	
		(0001)	$(01\bar{1}0)$
Hyp-Pro-Gly	1	-429	-666
	2	-421	-654
	3	-407	-637
	4	-403	-555
Pro-Hyp-Gly	1	-450	-625
	2	-410	-599
	3	-378	-536
	4	-362	-508
Pro-Lys-Gly	1	-657	-756
	2	-599	-719
	3	-537	-589
	4	-371	-485
Pro-Hyl-Gly	1	-694	-872
	2	-644	-640
	3	-621	-596
	4	-466	-494

As mentioned, the nature of the side chain of the amino acid residues that form the peptides is one of the major factors determining the adsorption energy. When we compare the adsorption energies for the different peptides, we see that the presence of a basic residue in the chain of the PRO-LYS-GLY and PRO-HYL-

GLY peptides favours a strong electrostatic interaction between the peptide and hydroxyapatite. As a result, the energy released upon adsorption of PRO-LYS-GLY and PRO-HYL-GLY are larger than for HYP-PRO-GLY and PRO-HYP-GLY, due to the positively charged NH_3^+ functional group of the lysine residue. The hydroxylation of the lysine residue creates further binding opportunities to calcium and phosphate ions, and sorption of PRO-HYL-GLY peptide is hence energetically the most favourable.

The position of the hydroxyproline residue in the HYP-PRO-GLY and PRO-HYP-GLY peptides does not affect the adsorption energies significantly, as the geometry of both peptides is very similar and the adsorption energies depend on the number of interactions that can be formed to the surface species. We see in Table 6.2 that in the case of the (0001) surface, PRO-HYP-GLY forms a greater number of close interactions between the carboxylate group and surface calcium ions, while HYP-PRO-GLY has more interactions to the $(01\bar{1}0)$ surface.

Table 6.2 HA-Peptide interatomic distances for energetically preferred configurations (Å)

Peptide	HA Surface	Ca···OOC-	Ca···O=C-/Ca···OH-C	Ca···N-	OH···O_{pep}	HO···HN	PO_4···H_{pep}
Hyp-Pro-Gly	(0001)	2.41, 2.61, 2.63	2.40	---	---	---	1.41, 2.43
	$(01\bar{1}0)$	2.38, 2.39, 2.88	2.31	2.70	---	---	2.30
Pro-Hyp-Gly	(0001)	2.31, 2.81, 2.81, 2.69	2.55/2.46	---	2.96	---	1.41, 2.80, 2.32, 2.38, 2.99
	$(01\bar{1}0)$	2.39, 2.39	2.53	2.51	2.62	2.85	2.77
Pro-Lys-Gly	(0001)	2.46, 2.49, 2.54, 2.65	2.49, 2.80	2.49	---	---	1.52, 1.53, 2.03, 2.80, 2.91, 2.94
	$(01\bar{1}0)$	2.38, 2.34, 2.52	2.45 2.62	2.69	---	---	1.76, 1.98, 2.12, 2.64
Pro-Hyl-Gly	(0001)	2.41, 2.47, 2.60, 2.61	2.61/2.40	---	---	---	1.59, 1.69, 1.93, 2.15
	$(01\bar{1}0)$	2.43, 2.46, 2.53	2.58/2.51	2.60	---	---	1.51, 2.14, 2.14, 2.57

6.3 Chapter Conclusions

In this chapter, we have presented a detailed simulation study of the adsorption mechanism of four peptides with different functional residues, HYP-PRO-GLY, PRO-HYP-GLY, PRO-LYS-GLY, and PRO-HYL-GLY to the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces. All four peptides adsorb more strongly to the less

stable $(01\bar{1}0)$ surface than the (0001) surface, with proton transfer generally occurring from the peptides to the reactive $(01\bar{1}0)$ surface. On the (0001) surface, only when the amino acid residues have charged polar side group (PRO-LYS-GLY and PRO-HYL-GLY) does the proton of the amine group of the lysine and hydroxylysine migrate to the basic phosphate group. Moreover, in this surface the peptides are adsorbed as neutral zwitterions or in their cationic form, whereas on the $(01\bar{1}0)$ surface, the peptides are generally adsorbed in their anionic form or as a neutral zwitterion.

All peptides form multiple interactions with the surface species, but the $–COO^-$ group was found to bind most closely with the calcium ions, although the $>NH_2^+$ interacts with the hydroxyapatite surfaces as well. However these groups are involved in peptide bonds in the chains of the collagen molecule and are therefore not free to interact with the surface, indicating that the functional side groups of amino acid residues should play a more important role in the interaction with the hydroxyapatite surfaces. Electrostatic attractions were generally observed between the surface and the positively charged side group ($–NH_3^+$) of the lysine and hydroxylysine peptides (PRO-LYS-GLY and the PRO-HYL-GLY), and the hydroxylation of lysine favoured the interaction between the PRO-HYL-GLY peptide and the surfaces even more. In the case of the PRO-HYP-GLY peptide, relatively stable hydrogen- bonds were formed between uncharged polar side groups (-OH), while in HYP-PRO-GLY the hydrophobic groups move away from the surface.

As the adsorption of collagen at hydroxyapatite surfaces takes place in an aqueous environment the presence of water molecules may affect the strength of adsorption and hence hydroxyapatite growth. The next chapter includes classical molecular dynamics simulations of the adsorption of peptides at hydroxyapatite surfaces in an aqueous environment. Although these methods cannot model chemical processes such as proton transfer, they can provide insight into hydration effects and the influence of temperature on the strength of adsorption.

Chapter 7 Modelling the interaction of a collagen peptide with hydroxyapatite surfaces in aqueous environment

In this chapter, classical molecular dynamics simulations have been used to investigate the adsorption of the peptide HYP-PRO-GLY on hydroxyapatite surfaces. The interatomic potential used to describe the interaction between peptide and hydroxyapatite has been validated with respect to density functional theory calculations. The binding process of HYP-PRO-GLY on the (0001) and $(01\bar{1}0)$ surfaces of hydroxyapatite has then been modelled in an aqueous environment, through the addition of explicit water molecules in the simulated system.

7.1 Introduction

Since water is the largest single component of the human body, when studying biochemical processes it is important to take into account the role played by water molecules. The adsorption of peptides to the dry surfaces of hydroxyapatite, as computed in chapter 6, is thus, not an entirely accurate simulation model of the actual process of the peptide adsorption on the hydroxyapatite surfaces. In this chapter, we consider therefore the adsorption of the peptide HYP-PRO-GLY on the (0001) and $(01\bar{1}0)$ surface of hydroxyapatite in presence of explicit water molecules. The simulated system consisted of an aqueous solution of density 0.99 g mL^{-1}, which represents a more realistic model for studying the interaction of the peptide with hydroxyapatite.

As already described in chapter 6, the peptide chain was constructed by linking the amino acid residues in the sequence hydroxyproline, proline and glycine ($^{+}H_2N$-Hydroxyproline-Proline- Glycine-COO^{-}), shown in Figure 7.1.

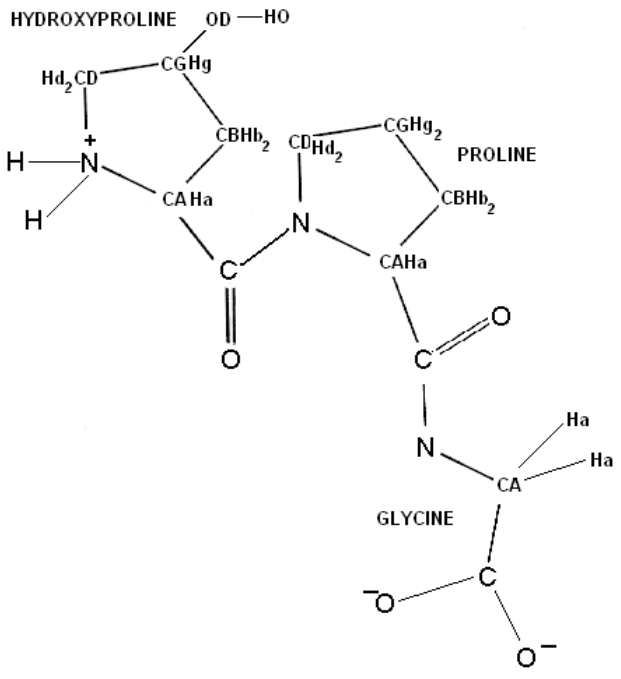


Figure 7.1 Schematic drawing of ($^{+}H_2N$-HYP-PRO-GLY-COO^{-}) peptide structure.

The adsorption energy of this peptide on hydroxyapatite slabs 20 Å thick was computed at DFT-GGA level as described in chapter 6. This was followed by molecular dynamics (MD) simulations, at constant volume and at temperature of 0*K, in vacuo.* This series of simulations allowed us to validate the accuracy of the interatomic potential model. Finally, MD simulations of the HYP-PRO-GLY/HA system were performed in aqueous environment at the conditions of constant pressure (p = 1 atm), constant temperature (T = 310 *K*) and with a solution density of 0.9922 g mL^{-1}.

7.2 Adsorption of HYP-PRO-GLY peptide at HA surfaces in vacuo: DFT Calculations

The surface simulation cells were supercells of sufficient size (262-354 $Å^2$), containing 528 atoms and a vacuum gap of at least 50 Å. If the cells had been too small the adsorbed peptide could have interfered with its images in the periodically repeated surface cell, which would affect the geometries and energies.

We have shown in section 6.2.1 that the peptide forms strong interactions with the hydroxyapatite surfaces, which confirms the affinity of the NH_3^+ and COO^- groups for HA, consistent with the estimates reported for different protein/HA systems (Chen *et al.*, 2008, Gibson *et al.*, 2006, Goobes *et al.*, 2007, Long *et al.*, 2001, Makrodimitris *et al.*, 2007, Shaw *et al.*, 2000, Capriotti *et al.*, 2007). The lowest energy geometry of the peptide adsorbed at the dehydrated (0001) surface is shown in Figure 7.2 (a, b) where one oxygen atom of the carboxylate group interacts simultaneously with two calcium ions (Ca···O = 2.41-2.61 Å), and the second oxygen of the carboxylate group interacts with a further calcium ion (Ca···O = 2.62 Å). One carbonyl group interacts with other surface calcium ions (Ca···O = 2.40 Å), while a range of hydrogen-bonds between the hydrogen atoms of the amine and hydroxy groups with oxygen atoms of the phosphate groups further enhance the binding between the peptide and the (0001) surface ($NH\cdots OPO_3$ = 1.4-2.7 Å, $-OH\cdots OPO_3$ = 2.4-2.7 Å).

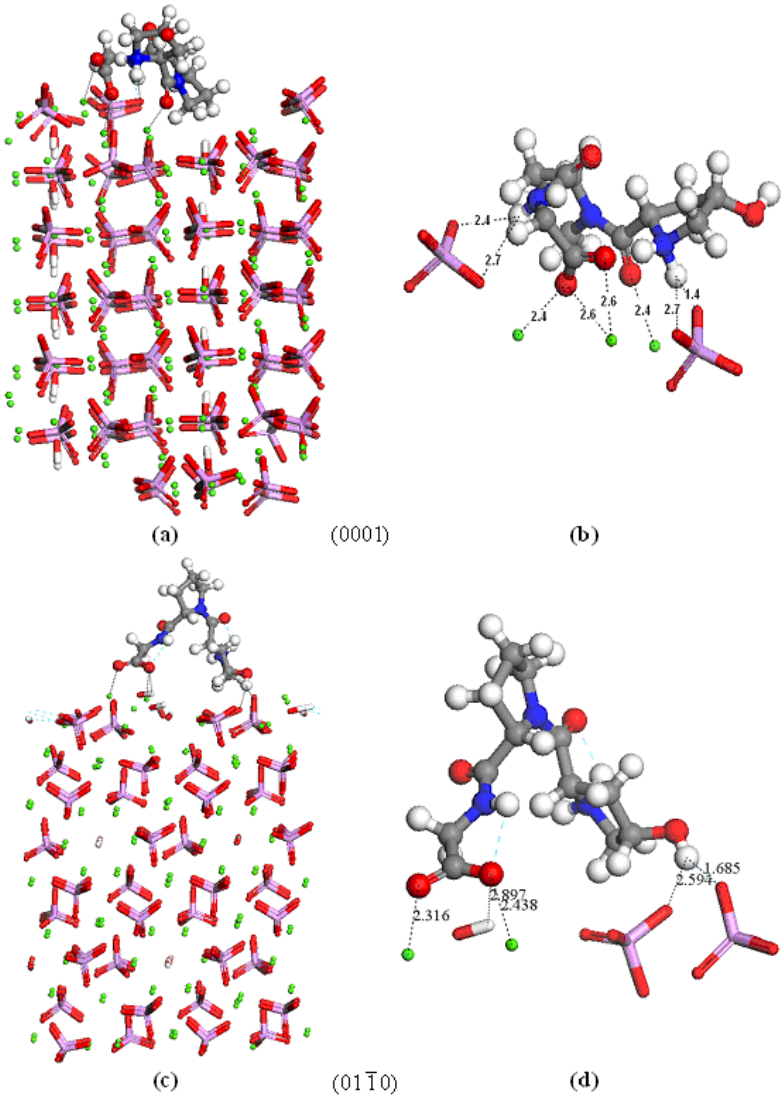


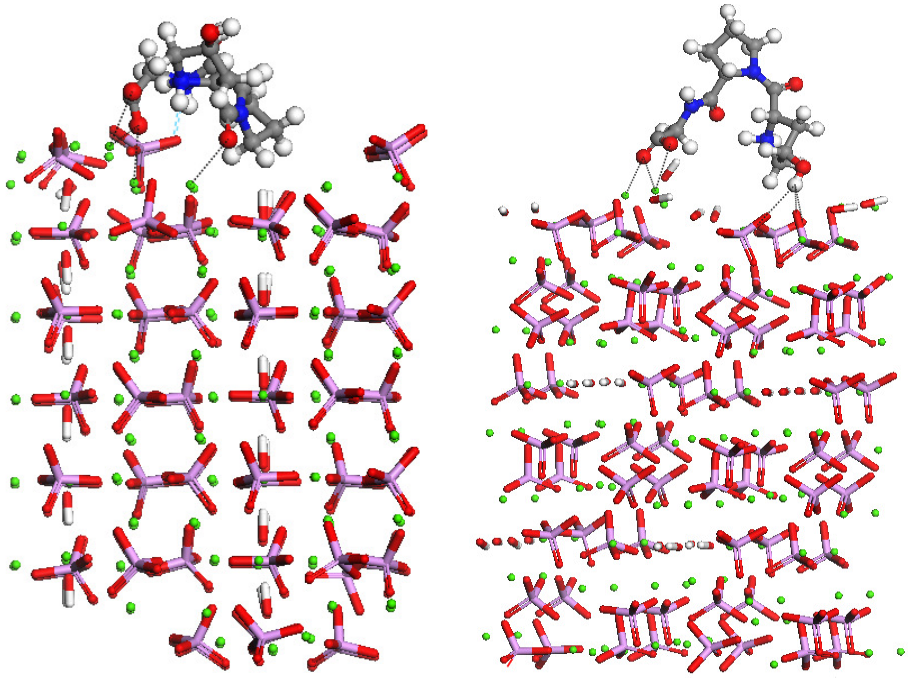
Figure 7.2 Geometry-optimised structures of the peptide adsorbed at the hydroxyapatite surfaces, showing interatomic distances: **(a)** overall and **(b)** detailed view of the (0001) surface, **(c)** overall and **(d)** detailed view of the $(01\bar{1}0)$ surface.

On the $(01\bar{1}0)$ surface, the strongest adsorption configuration includes chemical reactions (bond formation and breaking), such as proton transfer, which are not considered in classical molecular dynamics. For this reason, we have chosen the third configuration from chapter 6 to compare with our interactomic potential simulations, shown in Figure 7.2 (c, d). The oxygen atoms of the carboxylate

group are shared among two surface calcium ions (Ca···OCO^- = 2.3-2.4 Å), whereas hydrogen-bonding takes place between hydrogen atoms of the peptide's hydroxy group and oxygen atoms of a surface phosphate group (-OH···OPO_3 = 1.7-2.6 Å). In addition, the oxygen atoms of the carboxylate group form hydrogen-bonds with the hydrogen atom of the hydroxy group (COO···HO = 2.9 Å.

7.3 Adsorption of HYP-PRO-GLY peptide at HA surfaces in vacuo: MD Calculations

In order to compare the DFT calculations with interatomic potential-based simulations and hence evaluate the potential model for the surface-adsorbate interactions, we have also modelled the same system *in vacuo* and at 0 K with MD. The final structures of the surface/peptide complexes from the MD simulations are shown in Figure 7.3 (a) and (b), where the peptide clearly adsorbs in a similar way to the DFT results on both surfaces (see Figure 7.2 (a) and (b)). On the (0001) surface, we see that the peptide molecule adsorbs on the edge of the step on the surface, which is an under-coordinated and hence reactive site, although the molecule is too big to fit comfortably within the gap. The peptide forms a number of close interactions between the oxygen atoms of its carboxylate group and the topmost calcium ions (Ca···O = 2.62-2.68 Å), as well as a range of hydrogen bonded interactions through the hydrogen atoms of its amine group (NH···OPO_3 = 1.69 Å).



(a) **(b)**

Figure 7.3 Average structures of the peptide adsorbed at the hydroxyapatite surfaces **(a)** *in vacuo* at the (0001) surface **(b)** *in vacuo* at the $(01\bar{1}0)$ surface (Ca = green, O = red, P = purple, H = white, C = grey, N = blue).

On the $(01\bar{1}0)$ surface, the most favourable mode of adsorption for the peptide is through a number of close interactions between the molecule's oxygen atoms of the carboxylate group and two surface calcium ions (COO···Ca = 2.58-2.66 Å). In addition, hydrogen-bonding takes place between oxygen atoms of a surface phosphate group and the hydrogen atoms of the molecule's –OH group (-OH···OPO_3 = 2.61-2.79 Å). Further hydrogen-bonding occurs between the oxygen atoms of the carboxylate group and the hydrogen atoms of the two topmost hydroxy groups (COO···HO^- = 2.90 Å), whereas the non-polar proline residue does not form close interactions with the surface. Hydroxyapatite is a predominantly ionic crystal, where the phosphate and hydroxy groups act as poly-

anions (de Leeuw, 2001, 2002). It is therefore reasonable that the peptide's polar groups are largely responsible for its adsorption onto the ionic hydroxyapatite surfaces.

We have calculated the adsorption energies from both DFT calculations and MD simulations, which are listed in Table 7.1. When we compare the DFT and MD energies, the difference in adsorption energies between the two methods is of 65 kJ mol^{-1}, where the $(01\bar{1}0)$ surface is in both cases the more reactive towards adsorption of the peptide. In the MD simulations, the peptide is slightly less closely bound to the surface, which is reflected in the adsorption energies, but the agreement between the DFT calculations and classical MD simulations is excellent, especially taking into account the approximations inherent in both methods, the lack of vibrational properties in the DFT calculations, and the tendency of the generalised gradient approximation (GGA) used in our DFT calculations to lead to over-binding of surface adsorbates (Fernández *et al.*, 2007). More importantly, the relative adsorption energies for the two surfaces, calculated by both methods are exactly the same at 208 $kJmol^{-1}$, which is in excellent agreement. The difference in the energies between the two methods can be taken as an upper limit to any uncertainties in the adsorption energies.

Table 7.1 Calculated adsorption energies (kJ mol^{-1}) of the peptide at HA surfaces.

HA surface	DFT in *vacuo* (0K)	MD in *vacuo* (0K)
(0001)	-429	-365
$(01\bar{1}0)$	-637	-573

The potential model thus adequately models the hydroxyapatite crystal and the interactions of the peptide with the surfaces and we now introduce temperature and water into our MD simulations to model the system under more realistic conditions

7.4 Peptide in water

The solvated peptide was optimised both as a neutral and as an ionic molecule (zwitterion), which should be the preferred configuration in water. The lower-energy optimised structure of the peptide molecule is indeed the charge-separated configuration, where its stability is due to the formation of intra-molecular hydrogen-bonds. When we immersed the peptide in water, shown in Figure 7.4 (b), the geometry of the peptide changed compared to the vacuum geometry (Figure 7.4 (a)). The effect of the water has led to breaking and lengthening to some extent of most of the intra-molecular interactions in the molecule through the formation of inter-molecular hydrogen-bonded interactions with the surrounding water molecules. As expected, the water molecules interact through their hydrogen ions with the carboxylate, carbonyl, and hydroxy oxygens of the

peptide and through their oxygen ions with the hydrogen atoms of the amine groups of the peptide. The hydroxyproline and glycine residues in particular become solvated and form hydrogen-bonds with the water molecules, in agreement with MD simulations of collagen I in water (Handgraaf & Zerbetto, 2006).

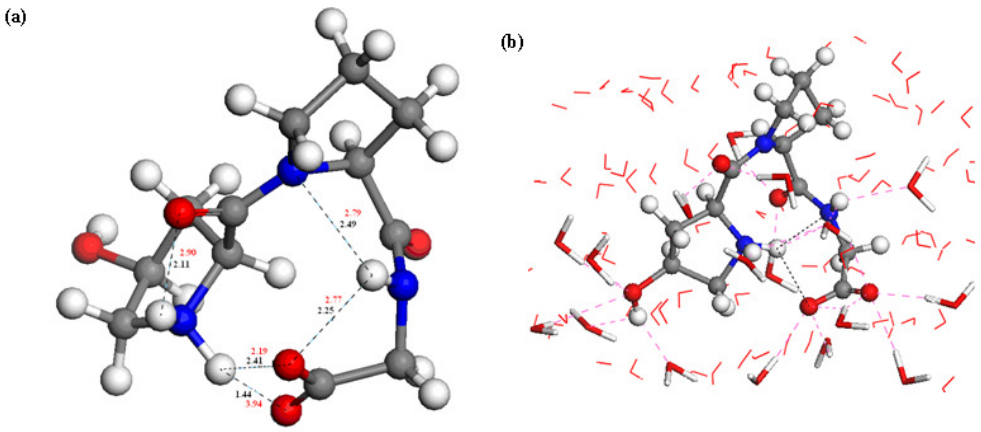


Figure 7.4 Optimised structure of ($^{+}H_2N$-HYP-PRO-GLY-COO^{-}) peptide structure in water, showing some of the extensive intra-molecular network of hydrogen bonding in black for the isolated molecule and in red when submerged in water (C = grey, O = red, N = blue, H = white).

7.5 Adsorption of the peptide in water

Figure 7.5 displays a series of representative snapshots of the peptide molecule adsorbed at the interface with the (0001) surface in water. In the initial configuration the peptide is on the edge of the step on the surface, as was seen in vacuo (Figure 7.5 (a)). Within the first two hundred picoseconds it is clear that the presence of water molecules has affected the geometry and adsorption sites of the peptide, as result the peptide is adsorbed on the terrace rather than on the step

edge, and has fewer interactions between the oxygen atoms of the carboxylic acid and carbonyl groups and surface calcium ions. Moreover, the smaller water molecules, which are of course free to translate and rotate, accommodate themselves more easily within the gap in the surface, whereas there is considerable hydrogen bonding between the peptide and surrounding water molecules (Figure 7.5 (b)). At 800 ps the hydrogen atoms of the amine group form hydrogen-bonds with the oxygen atom of the phosphate group, whereas oxygen atom of the carboxylate group keeps interaction with calcium ion (Figure 7.5 (c)).

In the last time, the peptide remains near the surface, where a number of close interactions are formed between the molecule's oxygen atoms of the carboxylate groups and two surface calcium ions (Table 7.2). In addition, hydrogen-bonding takes place between oxygen atoms of a surface phosphate group and the hydrogen atoms of the molecule's –OH and $>NH_2$ groups (-OH···OPO_3 = 1.54-1.79 Å, NH···OPO_3 = 1.75 Å).

Table 7.2 Surface-peptide interactomic distances for energetically preferred configurations (Å)

HA Surface	Ca···OOC-	OH^-···O_{pep}	OH^-···N	OH^-···HN	PO_4···H_{pep}
(0001)	2.62, 2.58	---	---	---	1.39, 1.90, 2.68, 2.87
$(01\bar{1}0)$	2.77, 2.77, 2.65	2.57, 2.04, 2.74	2.83, 2.89	1.82, 1.69, 2.11	---

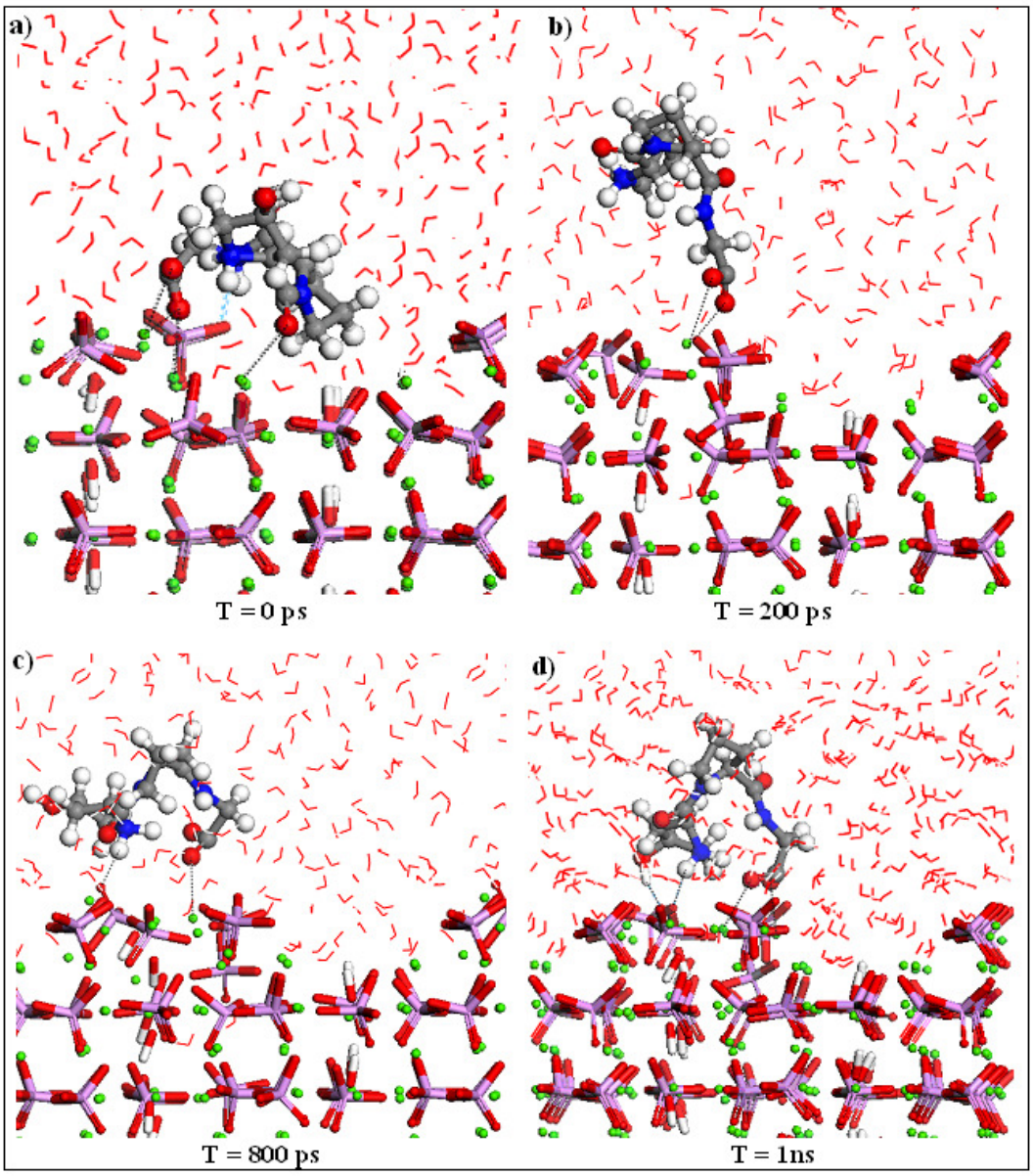


Figure 7.5 Representative snapshots from the MD simulation of the peptide adsorbed at the hydroxyapatite (0001) surfaces taken at T = 0 ps (a), 200 ps (b), 800 ps (c), and 1ns (f).

Figure 7.6 further reports on the dynamics of the peptide molecule adsorbed at the interface with the $(01\bar{1}0)$ surface in water, shown that there is no distinct preference for adsorption of water over surface. In all time of the simulation, the peptide prefers to remain near the surface than interact with surrounding water

molecules. On the other hand, water molecules replace the surface hydroxy groups in the surface, similar behaviour was found in the hydroxyapatite hydrated surface (see Chapter 4). The average adsorption mode for the peptide is through coordination by both carbonyl oxygen atoms and carboxylate group to two surface calcium atoms. One of the carbonyl oxygens is coordinated to a calcium at a distance of 2.77 Å, while the other interacts with another calcium at a distance of 2.46 Å, whereas a number of hydrogen-bonded interactions between its −OH and $>NH_2$ groups and oxygen atoms of the surface hydroxy groups further enhance the binding between the peptide molecule and the $(01\bar{1}0)$ surface (Table 7.2). Further hydrogen-bonding occurs between the oxygen atoms of the carboxylate and carbonyl groups and the hydrogen atoms of the two topmost hydroxy groups ($COO\cdots HO^-$ = 2.66 Å, $CO\cdots HO^-$ = 1.81 Å). Comparing Figure 7.5 and 7.6, it is clear that the presence of water has affected more geometry and adsorption site of the peptide on the (0001) surface than in the $(01\bar{1}0)$ surface.

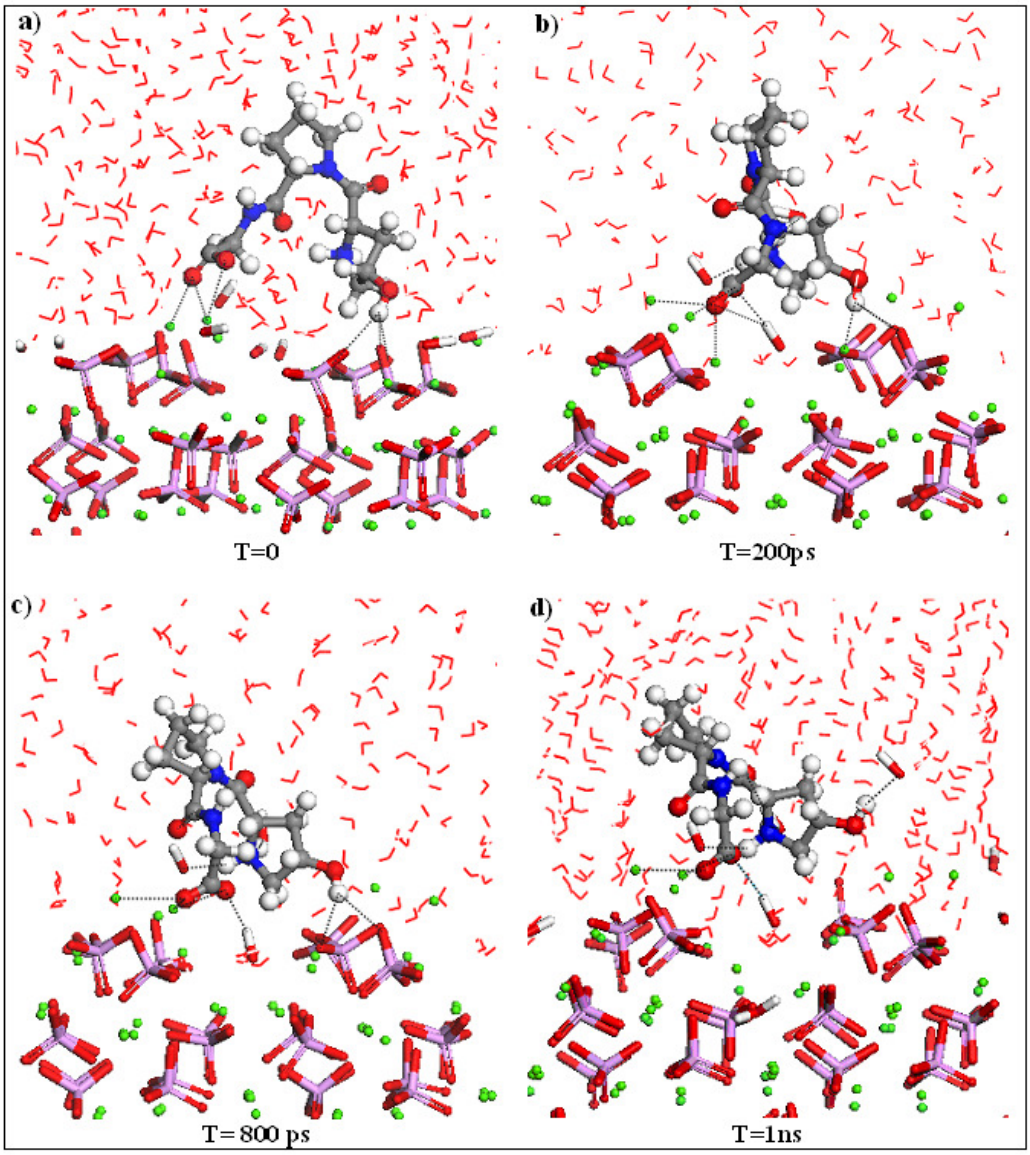


Figure 7.6 Representative snapshots from the MD simulation of the peptide adsorbed at the hydroxyapatite $(01\bar{1}0)$ surfaces taken at T = 0 ps (a), 200 ps (b), 800 ps (c), and 1ns (f).

7.6. Adsorption Energies

The adsorption energies of the peptide at the hydroxyapatite surfaces in aqueous environment, calculated as the difference in energy between the complete system and the sum of the energies of the hydrated surfaces and the solvated peptide, are

given in Table 7.2. The peptide interacts strongly with the two hydroxyapatite surfaces *in vacuo,* but in an aqueous environment the binding of the peptide to the (0001) surface has been destabilized by the presence, and indeed competition, of the water molecules. Whereas hydration of the $(01\bar{1}0)$ surface releases an energy of 314 kJmol^{-1} greater than hydration of the less reactive (0001) surface (see chapter 4), the difference in adsorption energies of the peptide at the hydroxyapatite surfaces increases from 133 kJ mol^{-1} in vacuo (at 310 K) to 520 kJ mol^{-1} in the aqueous environment, showing that the pre-adsorbed water has destabilised the peptide adsorption at the (0001) surface.

Table 7.3 Calculated adsorption energies (kJ mol^{-1}) of the peptide at hydroxyapatite surfaces

HA surface	MD in *vacuo* (310K)	MD in water (310K)
(0001)	-471	-157
$(01\bar{1}0)$	-604	-677

The large difference in absolute values between the adsorption energies *in vacuo* and in a hydrated environment at the (0001) surface, is due in part to the increased stability of the hydrated (0001) surface and the solvated peptide molecule, compared to the same systems *in vacuo*, where the peptide and the surface species could only increase their coordination through interaction with each other. In addition, on the (0001) surface the water adsorbs particularly strongly in the gap

in the surface, shifting the adsorption position of the peptide to the less reactive terrace, compared to the vacuum surface. As a result, adsorption of the peptide is no longer particularly favourable at the hydrated (0001) surface, but adsorption at the more reactive $(01\bar{1}0)$ surface is still highly exothermic and the peptide thus competes very effectively with the water molecules. Here, the geometry of the $(01\bar{1}0)$ surface does not lead to a shift in the preferred position of the peptide, which thus retains its strong interaction with this surface.

When adsorbing to the surfaces, the peptide also changes its conformation in order to match the surface geometry as closely as possible to maximise its interactions with the surfaces. We can obtain a quantitative measure of the conformational change on adsorption from a calculation of the change in conformational energy of the adsorbed peptide compared to the free peptide in its aqueous environment. The conformational energies of the peptide in their adsorbed geometries are calculated to be 35 $kJmol^{-1}$ higher at the (0001) surface and 238 $kJmol^{-1}$ higher at the $(01\bar{1}0)$ surface, compared to the solvated peptide conformation, indicating that the molecule's conformation is altered more upon adsorption at the $(01\bar{1}0)$ surface than at the (0001) surface. Despite this higher conformational energy, adsorption at the $(01\bar{1}0)$ surface is still far more energetically favourable than at the (0001) surface, indicating that the close interaction at the surface outweighs the gain in conformational energy.

The main interactions between the peptide and the surface are to surface Ca ions, and as is shown in the radial distribution function in Figure 7.7, the Ca-peptide interactions at the $(01\bar{1}0)$ surface are more pronounced than at the (0001) surface with Ca···O distances which are only marginally longer than fully ionic Ca···O bonds in solids (2.4 Å). Ionic Ca-O interactions between the surface and the peptide are energetically more favourable than the weaker hydrogen-bonded interactions between the solvated peptide to surrounding water molecules. Thus, if the peptide can form a number of Ca···O interactions at the surface, as on the $(01\bar{1}0)$ surface, this effect outweighs the increase in conformational energy required to accommodate the peptide to the surface geometry.

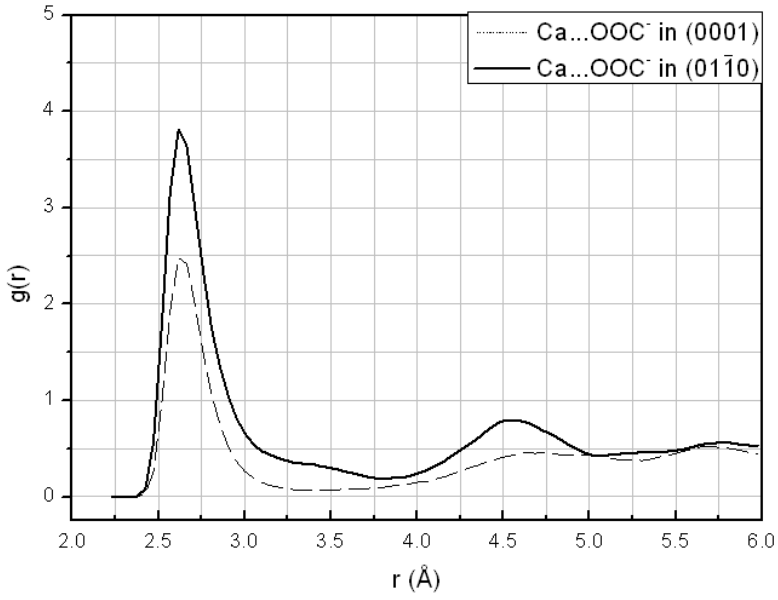


Figure 7.7 Radial distribution function between Ca-O_{pep} in hydroxyapatite hydrated surfaces (--- (0001) and — $(01\bar{1}0)$ surfaces).

7.7 Chapter Conclusions

In this chapter we have employed molecular dynamics simulations to investigate the interaction of a small peptide with the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces. To this end, we have used interatomic potential-based models, whose accuracy has been tested with respect to DFT calculations of the same systems, as to mode and strength of binding of the peptide at the surfaces.

We have executed a series of molecular dynamics simulations to compare the adsorption of the peptide at the hydroxyapatite surfaces in an aqueous environment, where our simulations show that the strength of interaction of the peptide with the surfaces depends on a number of factors:

- The stability of the surface, where the more reactive surface interacts more strongly with the adsorbing species, as shown by both DFT and MD calculations of the adsorption of the peptide at the vacuum surfaces, where the peptide was found to bind more strongly to the less stable $(01\bar{1}0)$ surface than the (0001) surface;
- The surface geometry, where the capability of the peptide molecule to form multiple interactions with the surface species is important, particularly if it can bridge between two or more surface calcium ions. However, any conformational changes in the peptide compared to its optimum geometry in water do not affect the binding strength, as the

strong interactions with the surface replace the weaker peptide-water interactions;

- The degree of competition from pre-adsorbed water molecules for surface adsorption sites, where we found that on occasion the smaller and more flexible water molecules effectively block the adsorption of the peptide at the most favourable surface sites.

As a result of a combination of these factors, the HYP-PRO-GLY peptide binds far more strongly to the hydroxyapatite $(01\bar{1}0)$ surface in an aqueous environment than to the thermodynamically preferred (0001) surface. Our study explains at the atomic-level, the experimental findings that only the $(01\bar{1}0)$ surface interacts strongly with peptides, which helps to explain the biological morphology of the hydroxyapatite mineral, where this surface is expressed preferentially as a result of the growth-directing effect of the protein matrix. Furthermore, the HYP-PRO-GLY peptide should be a good growth modifier for synthetic hydroxyapatite, leading to morphologies closer to the natural materials.

Chapter 8 Nucleation and Growth of Apatite at Collagen Templates. Preliminary Results

In the previous chapter we have employed classical molecular dynamics simulations to determine the mode out strength of peptide binding of the surface of hydroxyapatite. The interatomic potential used to describe the peptide/hydroxyapatite interaction gave results in good agreement with DFT calculations. In this chapter, we present therefore large-size and longer time-scale molecular dynamics simulations based on this interatomic potential to investigate the nucleation and growth of calcium phosphate clusters at the collagen template.

8.1 Introduction

Proteins are macromolecules that are formed by linear chains of amino acids linked by peptide bonds. This sequence of amino acids is called the primary structure of the protein, which is the most fundamental at the structural level. However, proteins are complex molecules, so there are a number of higher levels to describe their structure. The secondary structure in proteins consists of local inter-residue interactions mediated by non-bonding interaction, like hydrogen bonds. The most common secondary structures are the alpha helix and beta-pleated sheets. The tertiary structure of the protein refers to the way in which the entire protein molecule is coiled into an overall three dimensional shape, and the

quaternary structure refers to the way in which several protein molecules come together to yield large aggregate structures.

Proteins are classified as globular or fibrous, depending on their secondary and tertiary structures. Globular proteins such as hemoglobin (Figure 8.1 (a)) are water soluble, roughly spherical in shape, and mobile within cells. Fibrous proteins such as collagen (Figure 8.1 (b)) are tough, fairly rigid, and water insoluble and are used in nature for forming structures such as hair, nails and bones.

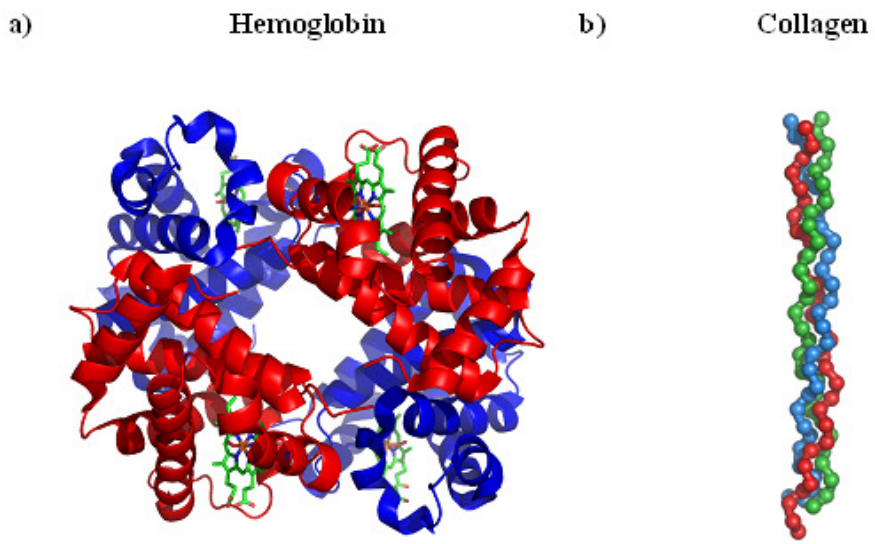


Figure 8.1 Structure of: **a)** haemoglobin, a globular protein and **b)** collagen, a fibrous protein

The collagen molecule, as described in chapter 1, is a triple-helical protein, rich in glycine and proline. The primary structure of the collagen is unusual: glycine accounts for one third of its amino acid content. Proline and hydroxyproline are frequently in the other position. In bone formation, the collagen assemble into

cross-triated fibrils, and a calcium phosphate, hydroxyapatite crystals, are arranged within and around the collagen fibrils.

Due to the importance of the apatite/collagen in both natural bone and in synthetic biomaterials, in this chapter, we have started to investigate the nucleation and growth of calcium phosphate clusters at the collagen template. In fact, the apatitic phase in bone is often formed from calcium phosphate precursor phases, such as octa-calcium phosphate, which is structurally similar to apatite in some regions and is thought to exist at the surface of the growing bone apatite before crystallising into the hydroxyapatite proper (Brown & Chow, 1976). As this nucleation/growth process obviously occurs in body fluid, we include water molecules in the calculations.

A hydrated crystal structure of collagen was obtained from the RSCB protein database (PDB ID: 1CGD). The structure was modified with Material Studio v 4.0 package, and contained three chains of $[^{-}OOC\text{-}(GLY\text{-}PRO\text{-}HYP)_{10}\text{-}NH_2^{+}]_3$ folded into a triple helix collagen molecule. The final system consisted of one $[^{-}OOC\text{-}(GLY\text{-}PRO\text{-}HYP)_{10}\text{-}NH_2^{+}]_3$ collagen molecule, 120 Ca^{2+} ions, 72 PO_4^{3-} ions, 24 OH^{-} ions, and 5225 water molecules placed in a box of size 100 x 35 x 35 Å (Figure 8.2). The density of the simulation box was of 0.9922 gmL^{-1} at 310 K temperature. The equilibration period was of 500 ps, in which all the atoms were free to move at constant volume and temperature (NVT). The temperature was fixed using the Nosé-Hoover thermostat, and the Verlet leap-frog algorithm was

used to integrate the equations of motion with a time step of 0.1fs. The analysis of results was performed on the basis of a 2 ns run at constant number of particles, pressure, and temperature (NPT).

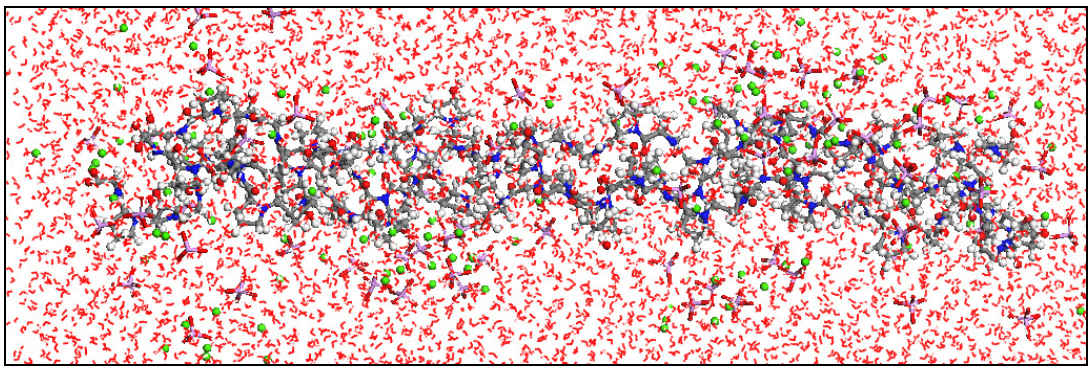


Figure 8.2 Graphical representation of the system: collagen shown in the center of the box as ball and stick model. Ca^{2+} ions as green ball, PO_4^{3-} and OH^- as stick model and line model of the water. Colours: H white, C gray, N blue, O red, and P purple.

In Figure 8.3, the water molecules have been removed to allow a clear view of the collagen molecule, Ca^{2+}, PO_4^{3-}, and OH^- ions at beginning of the equilibration period.

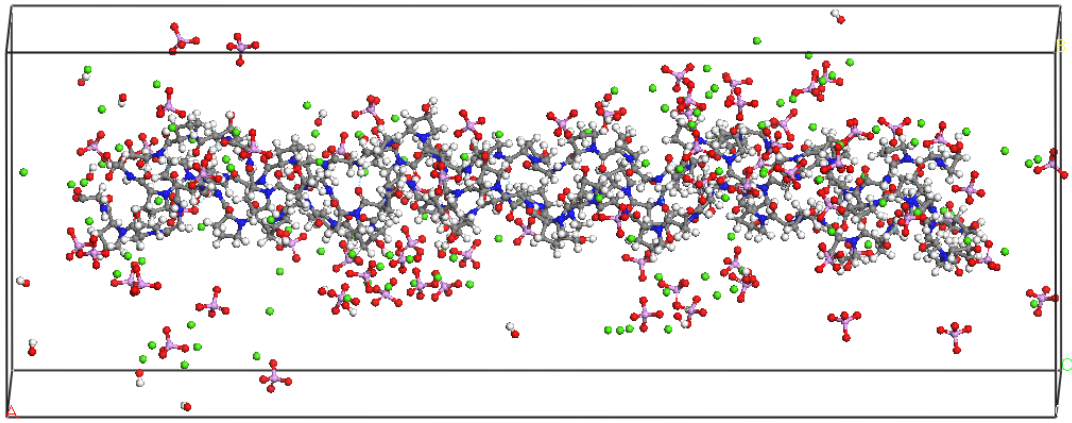


Figure 8.3 Graphical representation of the system: collagen shown in the center of the box as ball and stick model, Ca^{2+} ions as green ball, PO_4^{3-} and OH^- as stick model, having removed the water molecules.

8.2 NVT Equilibration Simulation at 310K

The first part of the simulation was performed in the NVT ensemble at body temperature (310 *K*). During the equilibration period, the hydration spheres are formed around the Ca^{2+} and PO_4^{3-} ions. Figure 8.4(a) shows how a number of water molecules have moved closer to the calcium ion. For Ca^{2+} the first maximum in radial distribution function (RDF) is at 2.38 Å (Figure 8.5), which corresponds to the positions of the oxygen atoms in the first coordination shell. The second coordination shell is less clearly defined since its associated peak is smaller and it spreads over a broader region, from 4.1 to 4.7 Å. The average number of the water molecules in the first hydration shell is between 6 and 8, where the numerical value of the coordination number probably depends on the distance from the oxygen atom of the phosphate group, which is in very good agreement with the values obtained from *ab initio* MD simulations on water by Di Tommaso *et al*. (Di Tommaso & de Leeuw, 2008).

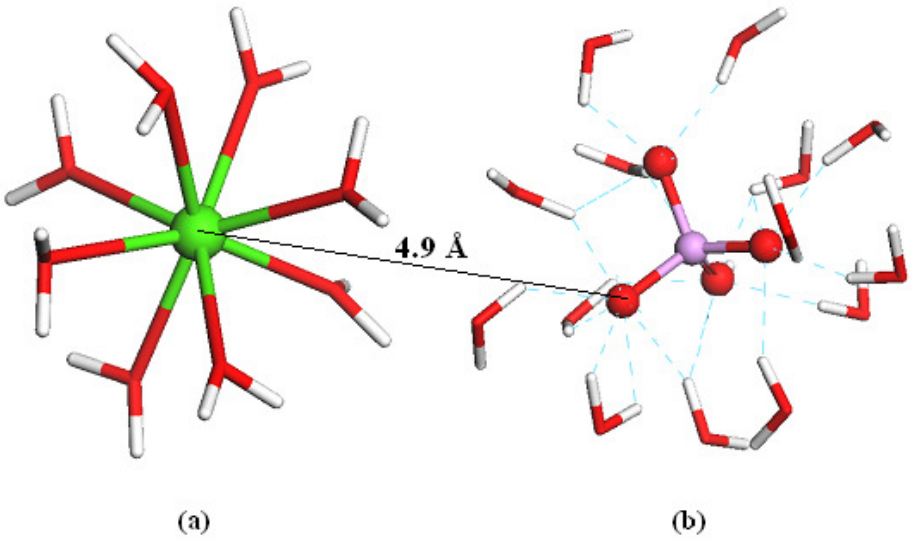


Figure 8.4 Snapshot of the Ca^{2+}/PO_4^{-3} ion pair with their surrounding water molecules. The ion pairs are formed by the $Ca(H_2O)_8^{2+}$ and the $PO_4(H_2O)_{14}^{-3}$ complexes and they move through the liquid with an average distance of 4.9 Å.

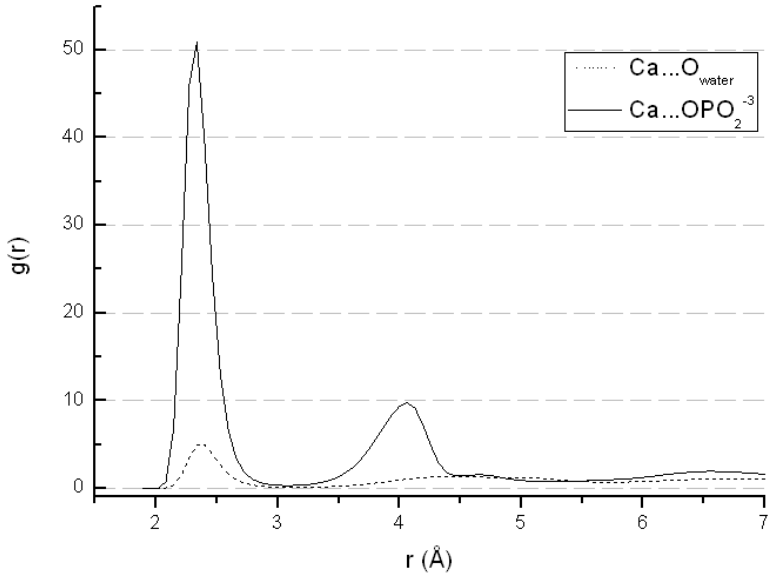


Figure 8.5 $Ca^{2+}\cdots O_{water}$ and $Ca^{2+}\cdots PO_4^{-3}$ Radial Distribution Functions.

Phosphate ions form stable complexes with the water molecules (Figure 8.4(b)). Figure 8.6 shows the $PO_4\cdots H_{water}$ and $PO_4\cdots O_{water}$ radial distribution function. The first sharp peak of the $PO_4\cdots O_{water}$ pair at 2.66 Å has a height of 2.03, with a second peak with a maximum of 1.34 at 4.36 Å. The RDF of the oxygen atoms of the phosphate with the hydrogen atoms of the surrounding water has the first maximum located at 1.88 Å, and a broader peak centred at about 4.08 Å. The integration of the numbers of all four RDF functions show a average value of about 12 water molecules in the first hydration shell. Our results are all in complete agreement with a neutron diffraction study of a K_3PO_4 aqueous solution (15 ± 3 water molecules) (Mason *et al.*, 2003), and with *ab initio* MD simulations of a single PO_4 ion in water (Tang *et al.*, 2009b).

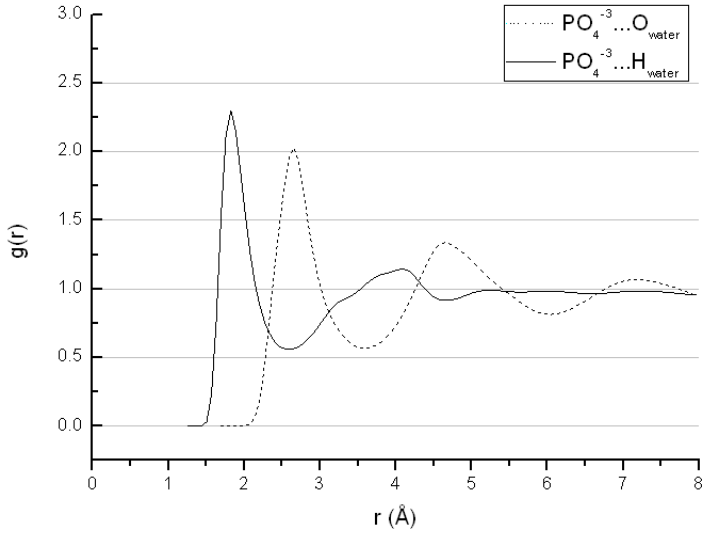


Figure 8.6. $PO_4^{-3}\cdots H_{water}$ and $PO_4^{-3}\cdots O_{water}$ Radial Distribution Functions.

Although the calcium and phosphate ions are stable in water, there is a strong electrostatic attraction between them, so it also found that calcium is coordinated with the phosphate atoms of oxygen at a distance of 2.33 Å as shown in the RDF, Figure 8.5.

These configurations in which the Ca^{2+} ion interacts with PO_4^{3-} anion are displayed in Figure 8.7. The ions diffuse through the liquid water, and form different types of ion pairs.

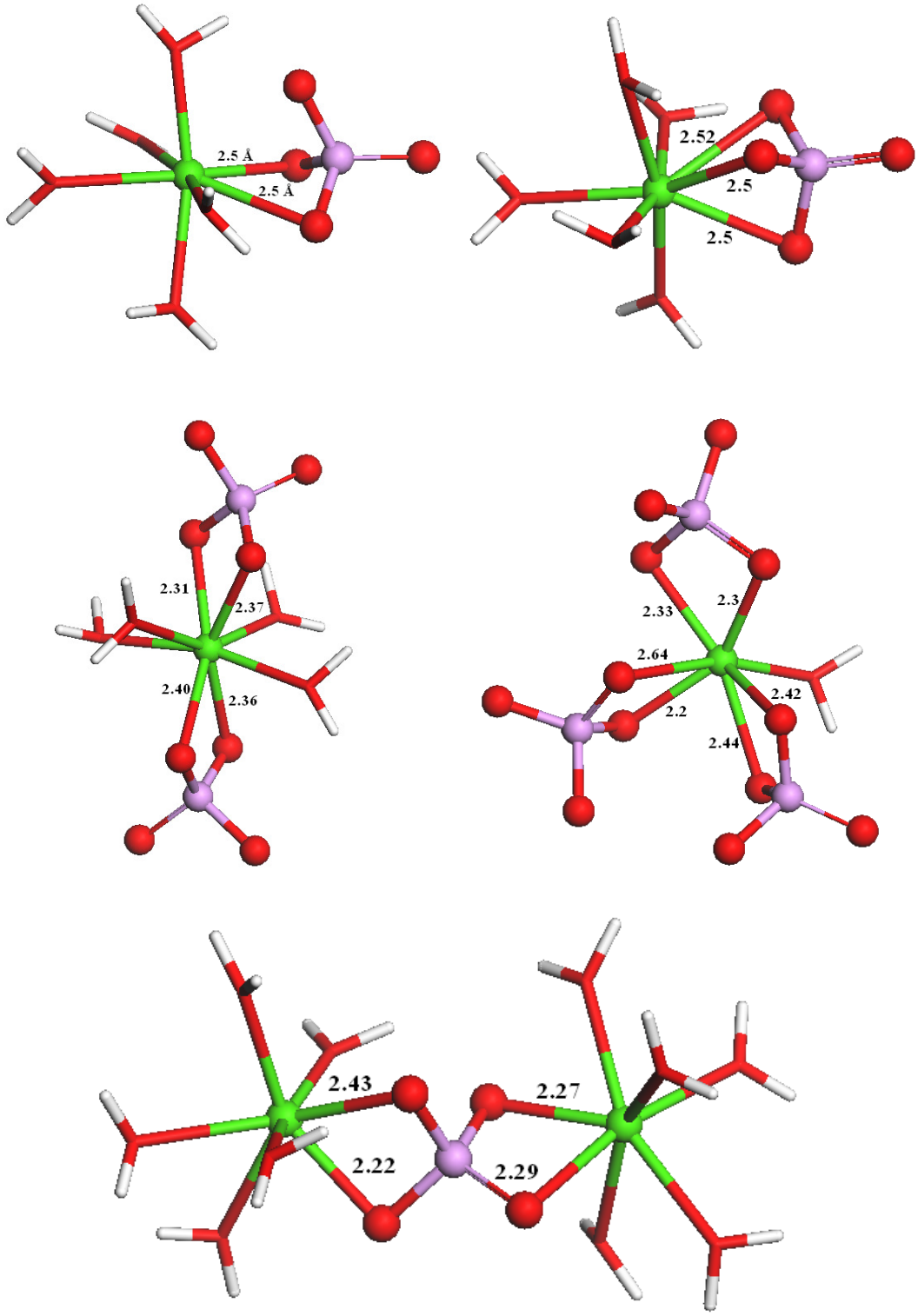


Figure 8.7 Structure of the five first clusters formed in solution, $Ca(H_2O)_5PO_4$, $Ca(H_2O)_4(PO_4)_2$, $Ca(H_2O)(PO_4)_3$, $Ca_2(H_2O)_{10}PO_4$.

The strong interaction between the hydroxide ions (OH^-) and the water molecules is reflected by the peak shifts of $OH^-\cdots H_{water}$ and $OH^-\cdots O_{water}$ RDF (Figure 8.8). It should be noted that the intermolecular peaks of these RDFs generally move to shorter distances compared to pure water. In fact, the hydrogen of OH^- is only weakly hydrogen-bonded to other oxygen, whereas the oxygen of OH^- forms several hydrogen-bondeds.

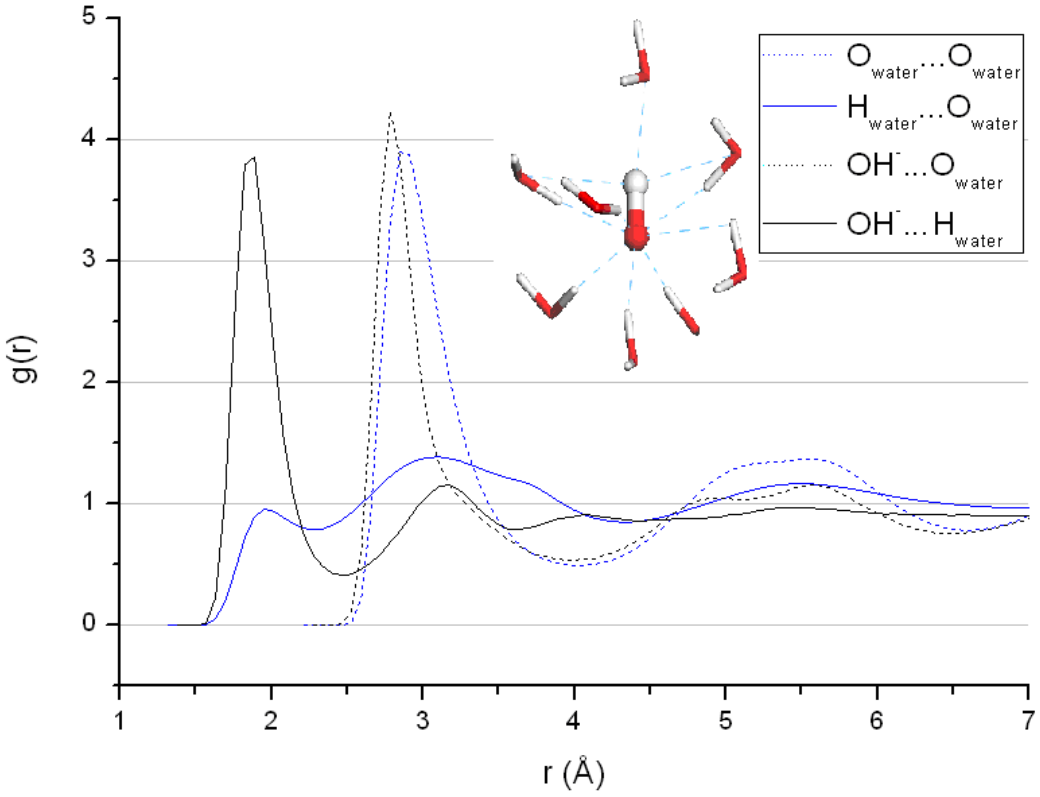


Figure 8.8 $OH^-\cdots H_{water}$ and $OH^-\cdots O_{water}$ radial distribution functions.

The radial distribution functions of water molecules with the collagen triple helix are shown in Figure 8.9. The water molecules form stronger hydrogen bonds with the hydroxy group of hydroxyproline residues, which interaction is more pronounced than with hydrogen atoms of amine group of the glycine residues or

oxygen atoms of any of the residues (1.94 Å). The second peak, only present in OH_{HYP}-water RDF, is 3.2 Å. The carbonyl oxygen of the proline is directed towards the center of the triple helix and forms hydrogen bonds with the amine hydrogen of the glycine in the neighbouring chain, leaving no space for water molecules.

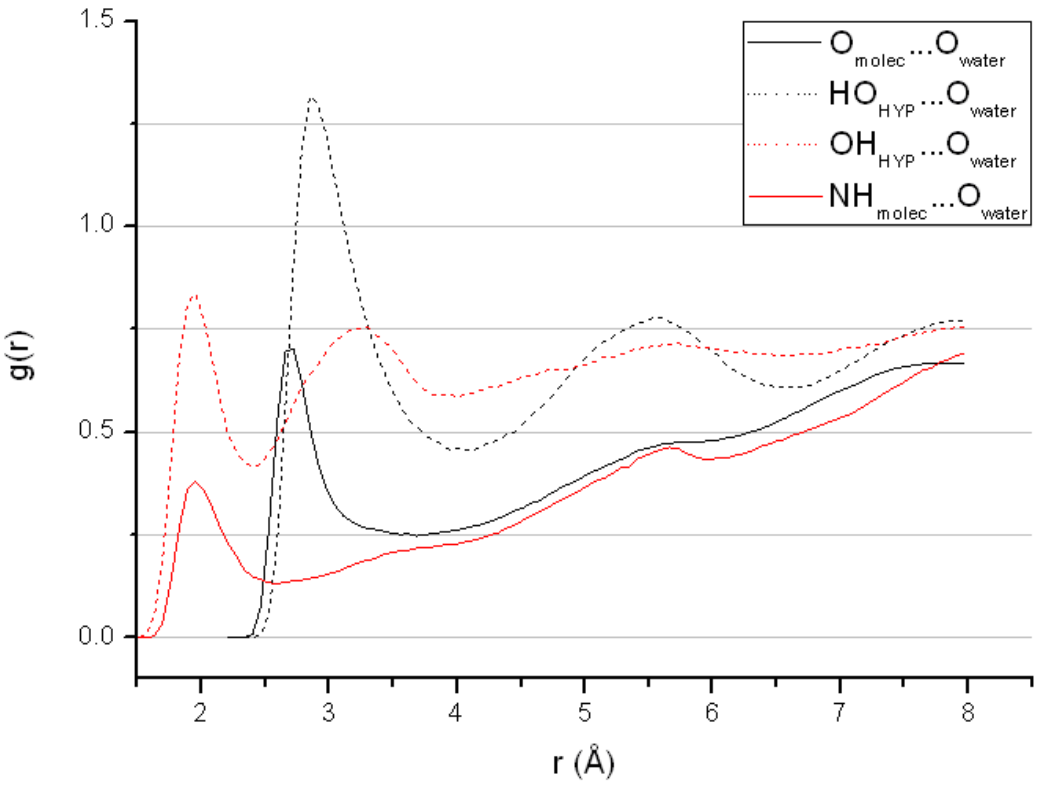


Figure 8.9 O_{molec}··· O_{water}, HO_{HYP}···O_{water}, HO_{HYP}···H_{water}, and NH_{molec}···H_{water} radial distribution functions.

We can see the distribution of the water molecules around the collagen in Figure 8.9, where the water molecules at a distance greater than 3 Å have been removed. Some of these characteristic hydrogen-bonding patterns of the water molecules have already been reported for the structures of model collagen peptides (Okuyama *et al.*, 2004, Handgraaf & Zerbetto, 2006, Kramer *et al.*, 2000, Kramer

et al., 2001, Nagarajan *et al.*, 1998, Kramer *et al.*, 1998, Vitagliano *et al.*, 2001, Bella *et al.*, 1995).

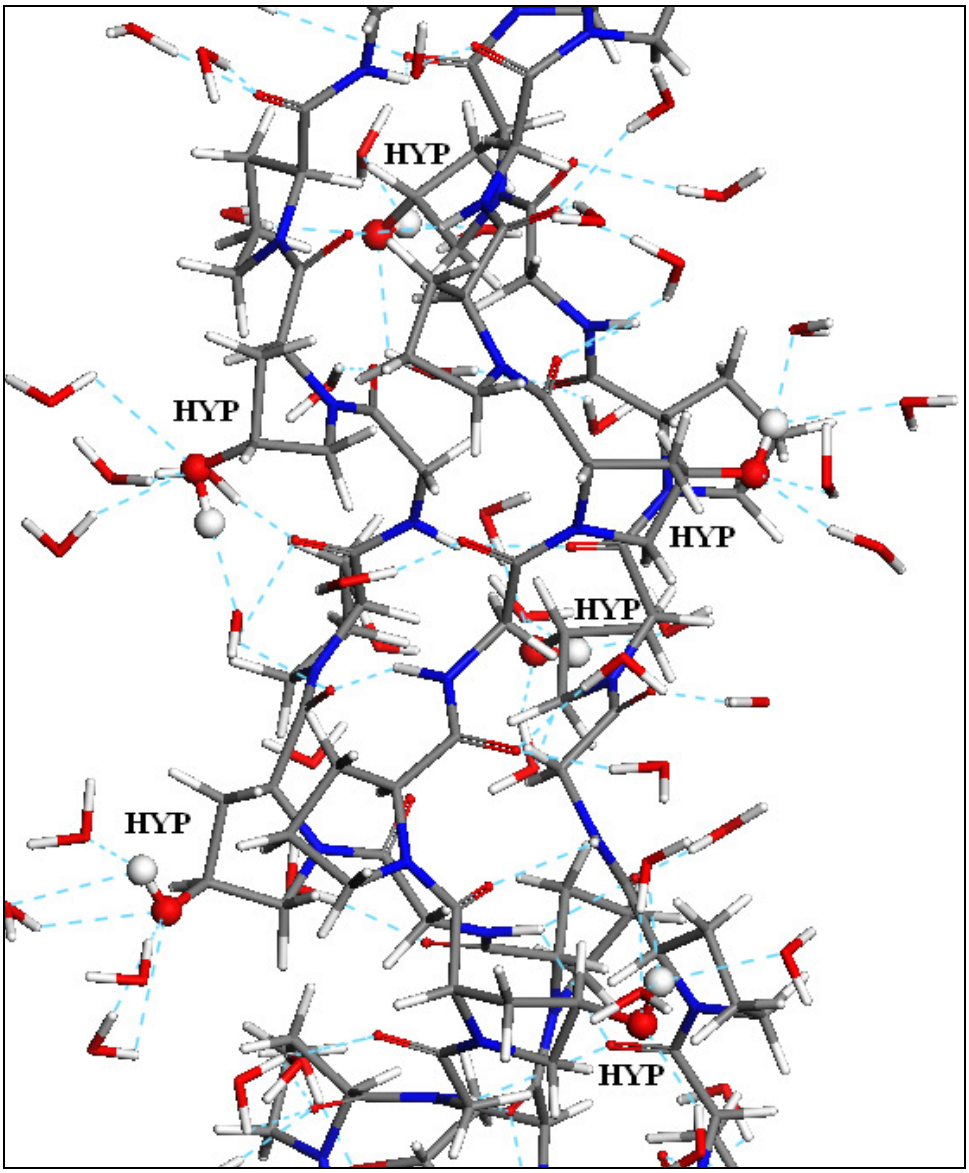


Figure 8.9 Graphical representation of the solvation shell of the collagen triple helix for an instantaneous configuration of the MD simulation. —OH of the HYP residues are shown in ball and stick model, other residues as sticks and water molecules at a distance of 1 to 3 Å from the collagen are indicated in blue.

8.3 NPT Simulation at 310K

The NVT simulations described in the previous section were used as an equilibration period. We now proceed to give the detailed results obtained from MD simulations in NPT ensemble during the first 2 ns, which is a production run without scaling of the temperature.

Figure 8.10 compares the RDFs for the Ca···O_{water}, Ca···PO_4 and Ca···$O_{collagen}$ pairs obtained during the equilibration period (t = 0 ps) after 2 ns of NPT simulations. After the equilibration period, the peak associated with the interactions between Ca and PO_4 is already formed (Figure 8.10 (b)). In the first 1.0 ns the first peak becomes higher and sharper, indicating that more Ca···PO_4 bonds are formed and the first aggregates are formed, whereas the number of water molecules in the hydration spheres around the calcium ions become fewer as time advances, as shown by the height of the first peak in the graph of the radial distribution function (Figure 8.10 (a)). The interactions between Ca and oxygen atoms of carboxylate, carbonyl and hydroxy groups increase during the simulation (Figure 8.10 (c)), as the first and second peak at 2.72 and 4.6 Å becoming more intense.

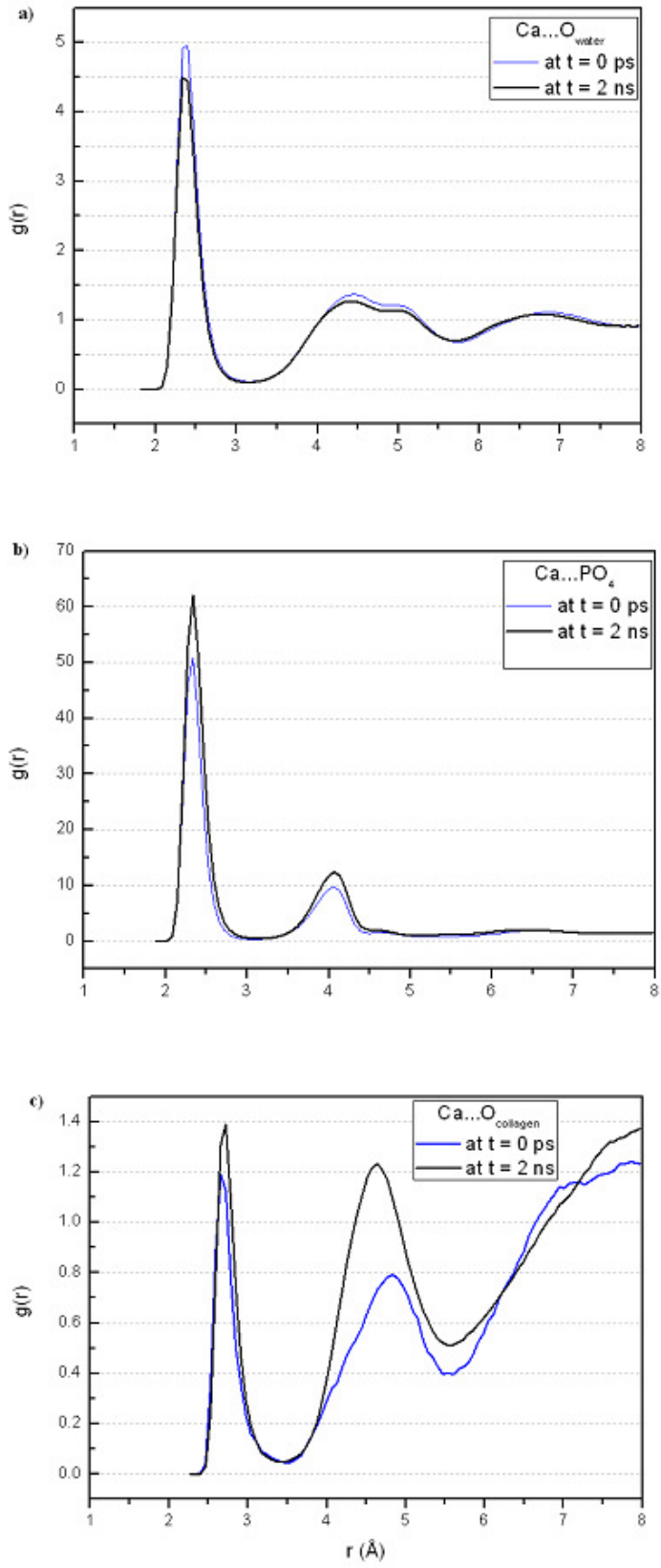


Figure 8.10 Ca···O_{water}, Ca···PO_4^{-3} and Ca···$O_{collagen}$ radial distribution functions at t = 0 and 2 ns.

Finally, it is interesting to consider in Figure 8.11 the variation of Ca···PO_4 carbonate distance as a function of time for the oxygens of the phosphate ion. We note that during the first picoseconds, the clusters start to grow. After 1 ns, clusters are formed in which a PO_4 group is surrounded by three Ca atoms placed at the vertices of a triangle.

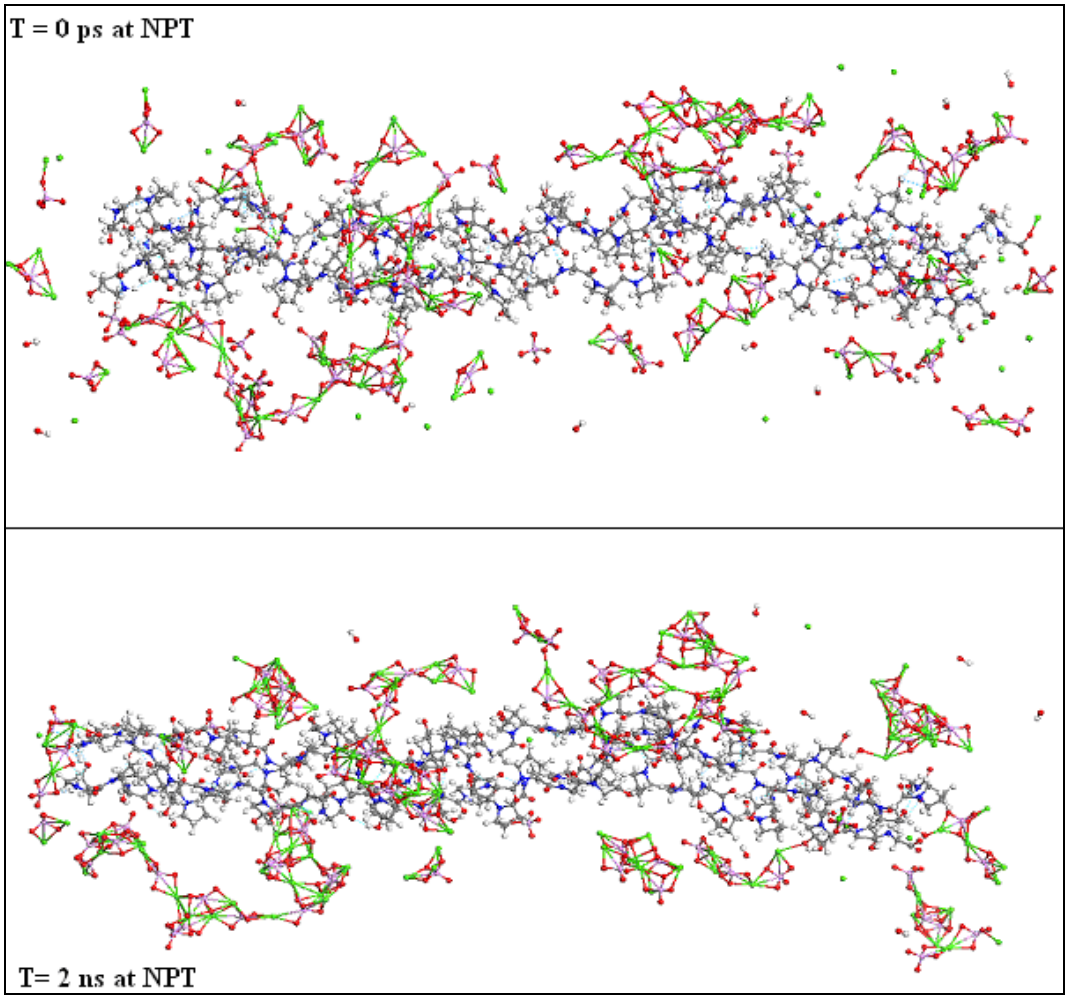


Figure 8.11 Snapshot of the system at after equilibration time and 2 ns. Water molecules have been removed.

In contrast to Kawska *et al.*(Tlatlik *et al.*, 2006, Kawska *et al.*, 2008), we have not found that the calcium ions may be incorporated between the strands of the triple

helix; the calcium ions coordinate with oxygens of the phosphate groups, whereas the hydroxy groups keep in solution. The initial formation of clusters of calcium phosphate in our calculation is in agreement with experimental studies, which report the initial formation of the more soluble octacalcium phosphate precursor phase for hydroxyapatite (Tarasevich *et al.*, 2007).

As many previous studies have found, the process of nucleation and crystal growth is not simple (Tarasevich *et al.*, 2007, Zahn *et al.*, 2007). The simulations are very time-consuming, with long calculations needed to simulate relatively short real times. For this reason, we only have preliminary results at this stage. However, the observed association of calcium, phosphate and hydroxy ions to a collagen protein reflects the very initial step of hydroxyapatite nucleation at a collagen matrix.

8.4 Summary

It is clearly a long way to go from the association of single ions into clusters to the formation of apatite/collagen composites, and our molecular dynamics simulations should be considered to model the initial stages. For the association of the calcium ions, the carboxylate group in the ends of the peptide strands were found to be freely accessible to the ions. Moreover, the hydroxy and carbonyl oxygen atoms may also support calcium collagen bonds. In the case of phosphate ions less favourable adsorption sites were observed. However, relatively stable

hydrogen-bonding interactions were found between oxygen atoms of the phosphate groups and hydrogen atoms of —OH of the hydroxyproline residues.

Once a Ca···PO_4 cluster is formed, it is relatively easy for it to continue growing in the presence of collagen triple helix. Of course, species present in the blood plasma such as extracellular proteins, and glycoproteins will also have roles in the nucleation processes and may speed up the growth of the hydroxyapatite.

Conclusions and Future work

8.1 Conclusions

In this thesis, we have presented a computer simulation study of the interactions of amino acids, peptides and a collagen molecule with hydroxyapatite, using both *ab initio* quantum-mechanical methods based on the density functional theory and interatomic potential simulations.

We have investigated the structure of hydroxyapatite, firstly, with electronic structure calculations based on the density functional theory to investigate the geometry of the perfect crystal, followed by calculations of the dry surfaces, where the calculated surface energies for the important experimental HA surfaces, showed the dominance of the dry (0001) surface over the and $(01\bar{1}0)$ plane. We then extended our study to include molecular dynamic calculations of the interaction of water with the surfaces. The simulations of the dry (0001) and $(01\bar{1}0)$ surfaces have identified that the stability of the (0001) surface is due to a combination of relative high coordination numbers of the surface species and large interlayer spacings. When hydrated, the $(01\bar{1}0)$ surface, which is highly reactive due to its large number of under-coordinated surface species, is stabilised but the (0001) surface remains by far the more stable surface, which agrees with experiment where the (0001) surface is expressed in the crystal morphology.

We have modelled the adsorption of the amino acids, glycine, proline and hydroxyproline onto the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces, where we found that all amino acid molecules coordinate to the surface via their carboxyl acid oxygen atoms to surface calcium atoms. All three amino acids are adsorbed as zwitterions at the hydroxyapatite (0001) surface, where the —COO^- group interacts with surface calcium ions and the —NH_3^+ (or >NH_2^+) hydrogen-bond to surface oxygen atoms. On the $(01\bar{1}0)$ surface, the acidic —COOH proton is transferred to the surface PO_4 group, leaving the amino acids negatively charged. Hydroxyproline is energetically the most favourable adsorbate, where the enhanced interaction of this molecule with the HA surfaces compared to proline, is due to the extra hydroxy functional group.

We next extended our study to include the adsorption of four peptides with different functional residues, HYP-PRO-GLY, PRO-HYP-GLY, PRO-LYS-GLY, and PRO-HYL-GLY to the hydroxyapatite (0001) and $(01\bar{1}0)$ surfaces. Again all four peptides adsorb more strongly to the less stable $(01\bar{1}0)$ surface than the (0001) surface, with proton transfer generally occurring from the peptides to the reactive $(01\bar{1}0)$ surface. On the (0001) surface, where only the amino acid residues have charged polar side groups (PRO-LYS-GLY and PRO-HYL-GLY) does the proton of the amine group of the lysine and hydroxylysine migrate to the basic phosphate group. In this surface, the peptides are adsorbed as neutral zwitterions or in their cationic forms, whereas on the $(01\bar{1}0)$ surface, the peptides are generally adsorbed in their anionic form or as neutral zwitterions.

Thus, the process of adsorption of all different amino acids or peptides is energetically more favourable to the $(01\bar{1}0)$ surface than to the (0001) surface, where the molecules adsorb as isolated molecules to the dry surfaces.

We have also coadsorbed one of the peptides, namely HYP-PRO-GLY, with water to the (0001) and $(01\bar{1}0)$ surfaces to study the explicit effect of the water on the adsorption selectivity of the peptide. The inclusion of water in the calculations now showed a very clear preference for adsorption to the $(01\bar{1}0)$ surface than on to the thermodynamically preferred (0001) surface. Our study explains at the atomic-level, the experimental findings that only the $(01\bar{1}0)$ surface interacts strongly with peptides, which helps to explain the biological morphology of the hydroxyapatite mineral, where this surface is expressed preferentially as a result of the growth-directing effect of the protein matrix. Furthermore, the HYP-PRO-GLY peptide should be a good growth modifier for synthetic hydroxyapatite, leading to morphologies closer to the natural materials.

Finally, we have studied the initial stages of the nucleation of the hydroxyapatite mineral at a collagen molecule in aqueous solution. We observed that $Ca\cdots PO_4$ clusters are formed, which can continue to grow in the presence of the collagen triple helix. However, as the computational efforts of our studies of aggregate growth up to small clusters took several CPU years, the observation of full ordering in terms of an apatite nano-crystal would require prohibitively long simulations.

8.2 Future work

Having investigated the interactions of the predominant molecules in the collagen protein with the apatite mineral, despite the important results achieved in this thesis, we think that our study of nucleation and growth of apatite at collagen molecules is only a start and there is still more work to be done to complete this study. Among the possible directions for future research is the simulation of a model of the complete composite materials, including the collagen fibrils and a number of apatites platelets. We could also include other extracellular proteins, glycoproteins, and relevant ions such as sodium and carbonates to simulate a realistic body fluid.

References

Agnihotri, A. & Siedlecki, C. A. (2004). Langmuir 20, 8846-8852.

Almora-Barrios, N., Austen, K. F. & de Leeuw, N. H. (2009). Langmuir 25, 5018-5025.

Altmann, J. A., Handy, N. C. & Ingamells, V. E. (1996). International Journal of Quantum Chemistry 57, 533-542.

Amjad, Z., Koutsoukos, P. G. & Nancollas, G. H. (1984). Journal of Colloid and Interface Science 101, 250-256.

Anglada, E., Soler, J. M., Junquera, J. & Artacho, E. (2002). Phys Rev B 66, -.

Austen, K. F., White, T. O. H., Marmier, A., Parker, S. C., Artacho, E. & Dove, M. T. (2008). Journal of Physics-Condensed Matter 20.

Bachelet, G. B., Hamann, D. R. & Schluter, M. (1982). Physical Review B 26, 4199-4228.

Balasubramanian, V., Grusin, N. K., Bucher, R. W., Turitto, V. T. & Slack, S. M. (1999). Journal of Biomedical Materials Research 44, 253-260.

Barralet, J., Best, S. & Bonfield, W. (1998). Journal of Biomedical Materials Research 41, 79-86.

Barry, J. J. A., Silva, M., Shakesheff, K. M., Howdle, S. M. & Alexander, M. R. (2005). Advanced Functional Materials 15, 1134-1140.

Becke, A. D. (1988). Physical Review A 38, 3098-3100.

Beeman, D. (1976). Journal of Computational Physics 20, 130-139.

Bella, J., Brodsky, B. & Berman, H. M. (1995). Structure 3, 893-906.

Bertinetti, L., Drouet, C., Combes, C., Rey, C., Tampieri, A., Coluccia, S. & Martra, G. (2009). Langmuir 25, 5647-5654.

Boccaccini, A. R., Blaker, J. J., Maquet, V., Day, R. M. & Jerome, R. (2005). Materials Science & Engineering C-Biomimetic and Supramolecular Systems 25, 23-31.

Borisenko, K. B., Reavy, H. J., Zhao, Q. & Abel, E. W. (2008). Journal of Biomedical Materials Research Part A 86A, 1113-1121.

Born, M. & Huang, K. (1954). Dynamical theory of crystal lattices. Oxford: Clarendon Press.

Born, M. & Oppenheimer, R. (1927). Annalen Der Physik 84, 0457-0484.

Bres, E. F. & Hutchison, J. L. (2002). Journal of Biomedical Materials Research 63, 433-440.

Brown, W. E. & Chow, L. C. (1976). Annual Review of Materials Science 6, 213-236.

Cabanas, M. V., Rodriguez-Lorenzo, L. M. & Vallet-Regi, M. (2002). Chemistry of Materials 14, 3550-3555.

Cacciafesta, P., Humphris, A. D. L., Jandt, K. D. & Miles, M. J. (2000). Langmuir 16, 8167-8175.

Calderin, L., Stott, M. J. & Rubio, A. (2003). Physical Review B 67.

Capriotti, L. A., Beebe, T. P. & Schneider, J. P. (2007). Journal of the American Chemical Society 129, 5281-5287.

Chen, B., Ivanov, I., Park, J. M., Parrinello, M. & Klein, M. L. (2002). Journal of Physical Chemistry B 106, 12006-12016.

Chen, C. C., Boskey, A. L. & Rosenberg, L. C. (1984). Calcified Tissue International 36, 285-290.

Chen, P. H., Tseng, Y. H., Mou, Y., Tsai, Y. L., Guo, S. M., Huang, S. J., Yu, S. S. F. & Chan, J. C. C. (2008). Journal of the American Chemical Society 130, 2862-2868.

Chen, X., Wang, Q., Shen, J. W., Pan, H. H. & Wu, T. (2007). Journal of Physical Chemistry C 111, 1284-1290.

Chen, Y. L., Zhang, X. F., Gong, Y. D., Zhao, N. M., Zeng, T. Y. & Song, X. Q. (1999). Journal of Colloid and Interface Science 214, 38-45.

Cooper, T. G. & de Leeuw, N. H. (2004). Langmuir 20, 3984-3994.

Coquet, R., Hutchings, G. J., Taylor, S. H. & Willock, D. J. (2006). Journal of Materials Chemistry 16, 1978-1988.

Cormack, A. N., Lewis, R. J. & Goldstein, A. H. (2004). Journal of Physical Chemistry B 108, 20408-20418.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). Journal of the American Chemical Society 117, 5179-5197.

Corno, M., Busco, C., Bolis, V., Tosoni, S. & Ugliengo, P. (2009). Langmuir 25, 2188-2198.

Corno, M., Busco, C., Civalleri, B. & Ugliengo, P. (2006). Physical Chemistry Chemical Physics 8, 2464-2472.

Cruz, F., Lopes, J. N. C. & Calado, J. C. G. (2006a). Fluid Phase Equilibria 241, 51-58.

Cruz, F., Lopes, J. N. C. & Calado, J. C. G. (2006b). Journal of Physical Chemistry B 110, 4387-4392.

Cruz, F., Lopes, J. N. C., Calado, J. C. G. & da Piedade, M. E. M. (2005). Journal of Physical Chemistry B 109, 24473-24479.

Cusco, R., Guitian, F., de Aza, S. & Artus, L. (1998). Journal of the European Ceramic Society 18, 1301-1305.

Dalas, E. & Koutsoukos, P. G. (1989). Journal of the Chemical Society-Faraday Transactions I 85, 3159-3164.

Dalpi, M., Karayianni, E. & Koutsoukos, P. G. (1993). Journal of the Chemical Society-Faraday Transactions 89, 965-969.

de Leeuw, N. H. (2001). Chemical Communications 1646-1647.

de Leeuw, N. H. (2002). Phys Chem Chem Phys 4, 3865-3871.

de Leeuw, N. H. (2004a). J Phys Chem B 108, 1809-1811.

de Leeuw, N. H. (2004b). Physical Chemistry Chemical Physics 6, 1860-1866.

de Leeuw, N. H. & Cooper, T. G. (2004). Crystal Growth & Design 4, 123-133.

de Leeuw, N. H. & Mkhonto, D. (2003). Chemistry of Materials 15, 1567-1574.

de Leeuw, N. H. & Parker, S. C. (1998). Physical Review B-Condensed Matter 58, 13901-13908.

de Leeuw, N. H., Parker, S. C. & Rao, K. H. (1998). Langmuir 14, 5900-5906.

de Leeuw, N. H. & Rabone, J. A. L. (2007). Crystengcomm 9, 1178-1186.

Di Tommaso, D. & de Leeuw, N. H. (2008). Journal of Physical Chemistry B 112, 6965-6975.

Dick, B. G. & Overhauser, A. W. (1958). Physical Review 112, 90-103.

Ding, Y. B. & Kroghjespersen, K. (1992). Chem Phys Lett 199, 261-266.

Drouet, C., Carayon, M. T., Combes, C. & Rey, C. (2008). Materials Science & Engineering C-Biomimetic and Supramolecular Systems 28, 1544-1550.

Du, C., Falini, G., Fermani, S., Abbott, C. & Moradian-Oldak, J. (2005). Science 307, 1450-1454.

Du, Z. M. & de Leeuw, N. H. (2006). Dalton Transactions 2623-2634.

Duffy, D. M. & Harding, J. H. (2004a). Langmuir 20, 7630-7636.

Duffy, D. M. & Harding, J. H. (2004b). Langmuir 20, 7637-7642.

Eglin, D., Maalheem, S., Livage, J. & Coradin, T. (2006). Journal of Materials Science-Materials in Medicine 17, 161-167.

Elliott, J. C., Mackie, P. E. & Young, R. A. (1973). Science 180, 1055-1057.

Ellis, D. E., Terra, J., Warschkow, O., Jiang, M., Gonzalez, G. B., Okasinski, J. S., Bedzyk, M. J., Rossi, A. M. & Eon, J. G. (2006). Physical Chemistry Chemical Physics 8, 967-976.

Ewald, P. P. (1921). Annalen der Physik 64, 253-287.

Fernandez-Serra, M. V., Junquera, J., Jelsch, C., Lecomte, C. & Artacho, E. (2000). Solid State Commun 116, 395-400.

Fernández, E. M., Eglitis, R. I., Borstel, G. & Balbás, L. C. (2007). Computational Materials Science 39, 587-592.

Ferneyhough, R., Fincham, D., Price, G. D. & Gillan, M. J. (1994). Modelling and Simulation in Materials Science and Engineering 2, 1101-1110.

Fields, G. B., Lauer, J. L., Dori, Y., Forns, P., Yu, Y. C. & Tirrell, M. (1998). Biopolymers 47, 143-151.

Filgueiras, M. R. T., Mkhonto, D. & de Leeuw, N. H. (2006). J Cryst Growth 294, 60-68.

Filippi, C., Singh, D. J. & Umrigar, C. J. (1994). Physical Review B 50, 14947-14951.

Fratzl, P., Gupta, H. S., Paschalis, E. P. & Roschger, P. (2004). Journal of Materials Chemistry 14, 2115-2123.

Freeman, C. L., Harding, J. H., Cooke, D. J., Elliott, J. A., Lardge, J. S. & Duffy, D. M. (2007). Journal of Physical Chemistry C 111, 11943-11951.

Gao, H. J. & Kong, Y. (2004). Annual Review of Materials Research 34, 123-150.

Garcia, A., Elsasser, C., Zhu, J., Louie, S. G. & Cohen, M. L. (1992). Physical Review B 46, 9829-9832.

Gear, C. W. (1971). Numerical initial value problems in ordinary differential equations. Englewood Cliffs, N.J.,: Prentice-Hall.

George, A. & Veis, A. (2008). Chemical Reviews 108, 4670-4693.

Gibson, I. R., Best, S. M. & Bonfield, W. (1999). Journal of Biomedical Materials Research 44, 422-428.

Gibson, J. M., Popham, J. M., Raghunathan, V., Stayton, P. S. & Drobny, G. P. (2006). Journal of the American Chemical Society 128, 5364-5370.

Gill, P. E., Murray, W. & Wright, M. H. (1981). Practical Optimization. London: Adademic.

Goobes, R., Goobes, G., Shaw, W. J., Drobny, G. P., Campbell, C. T. & Stayton, P. S. (2007). Biochemistry 46, 4725-4733.

Grunkemeier, J. M. & Horbett, T. A. (1996). Journal of Molecular Recognition 9, 247-257.

Hajdu, J., Machin, P. A., Campbell, J. W., Greenhough, T. J., Clifton, I. J., Zurek, S., Gover, S., Johnson, L. N. & Elder, M. (1987). Nature 329, 178-181.

Hamann, D. R. (1996). Physical Review Letters 76, 660-663.

Hamann, D. R., Schluter, M. & Chiang, C. (1979). Physical Review Letters 43, 1494-1497.

Hamprecht, F. A., Cohen, A. J., Tozer, D. J. & Handy, N. C. (1998). Journal of Chemical Physics 109, 6264-6271.

Handgraaf, J. W. & Zerbetto, F. (2006). Proteins-Structure Function and Bioinformatics 64, 711-718.

Harding, J. H. & Duffy, D. M. (2006). Journal of Materials Chemistry 16, 1105-1112.

Hauptmann, S., Dufner, H., Brickmann, J., Kast, S. M. & Berry, R. S. (2003). Physical Chemistry Chemical Physics 5, 635-639.

Haverty, D., Tofail, S. A. M., Stanton, K. T. & McMonagle, J. B. (2005). Physical Review B 71.

Hidalgo, F., Sanchez-Castillo, A. & Noguez, C. (2009). Physical Review B 79.

Hirschfelder, J. O., Bird, R. B. & Curtiss, C. F. (1954). Molecular theory of gases and liquids. New York: Wiley.

Hockney, R. W. (1970). Journal Name: Methods Comput. Phys. 9: 135-211(1970).; Other Information: Orig. Receipt Date: 31-DEC-70 Medium: X.

Hohenberg, P. & Kohn, W. (1964). Physical Review B 136, B864-&.

Hoover, W. G. (1985). Physical Review A 31, 1695-1697.

Jack, K. S., Vizcarra, T. G. & Trau, M. (2007). Langmuir 23, 12233-12242.

Jha, L. J., Best, S. M., Knowles, J. C., Rehman, I., Santos, J. D. & Bonfield, W. (1997). Journal of Materials Science-Materials in Medicine 8, 185-191.

Johnson, G., Jenkins, M., McLean, K. M., Griesser, H. J., Kwak, J., Goodman, M. & Steele, J. G. (2000). Journal of Biomedical Materials Research 51, 612-624.

Jones, A., Slater, P. R. & Islam, M. S. (2008). Chemistry of Materials 20, 5055-5060.

Junquera, J., Paz, O., Sanchez-Portal, D. & Artacho, E. (2001). Phys Rev B 64, -.

KAY, M. (1964). CRYSTAL STRUCTURE OF HYDROXYAPATITE, Vol. 204, Nature, p. 1050.

Kerisit, S., Cooke, D. J., Spagnoli, D. & Parker, S. C. (2005). Journal of Materials Chemistry 15, 1454-1462.

Kirkham, J., Brookes, S. J., Shore, R. C., Wood, S. R., Smith, D. A., Zhang, J., Chen, H. F. & Robinson, C. (2002). Curr Opin Colloid In 7, 124-132.

Kleinman, L. & Bylander, D. M. (1982). Physical Review Letters 48, 1425-1428.

Kniep, R. & Simon, P. (2008). Angewandte Chemie-International Edition 47, 1405-1409.

Kohn, W. & Sham, L. J. (1965). Physical Review 140, 1133-&.

Kostyrko, T. & Krompiewski, S. (2008). Semiconductor Science and Technology 23.

Kotch, F. W. & Raines, R. T. (2006). Proceedings of the National Academy of Sciences of the United States of America 103, 3028-3033.

Koutsopoulos, S. & Dalas, E. (2000a). J Cryst Growth 216, 443-449.

Koutsopoulos, S. & Dalas, E. (2000b). Langmuir 16, 6739-6744.

Koutsopoulos, S. & Dalas, E. (2001). Langmuir 17, 1074-1079.

Koutsoukos, P. G. & Nancollas, G. H. (1981). Journal of Physical Chemistry 85, 2403-2408.

Koutsoukos, P. G. & Nancollas, G. H. (1986). Colloids and Surfaces 17, 361-370.

Kramer, R. Z., Bella, J., Brodsky, B. & Berman, H. M. (2001). Journal of Molecular Biology 311, 131-147.

Kramer, R. Z., Venugopal, M. G., Bella, J., Mayville, P., Brodsky, B. & Berman, H. M. (2000). Journal of Molecular Biology 301, 1191-1205.

Kramer, R. Z., Vitagliano, L., Bella, J., Berisio, R., Mazzarella, L., Brodsky, B., Zagari, A. & Berman, H. M. (1998). Journal of Molecular Biology 280, 623-638.

Lee, C. H., Singla, A. & Lee, Y. (2001). International Journal of Pharmaceutics 221, 1-22.

Lee, C. T., Yang, W. T. & Parr, R. G. (1988). Physical Review B 37, 785-789.

Lee, W. T., Dove, M. T. & Salje, E. K. H. (2000). Journal of Physics-Condensed Matter 12, 9829-9841.

Lindan, P. J. D. & Gillan, M. J. (1993). Journal of Physics-Condensed Matter 5, 1019-1030.

Long, J. R., Shaw, W. J., Stayton, P. S. & Drobny, G. P. (2001). Biochemistry 40, 15451-15455.

Lundqvist, M., Sethson, I. & Jonsson, B. H. (2005). Langmuir 21, 5974-5979.

Magne, D., Weiss, P., Bouler, J. M., Laboux, O. & Daculsi, G. (2001). Journal of Bone and Mineral Research 16, 750-757.

Makrodimitris, K., Masica, D. L., Kim, E. T. & Gray, J. J. (2007). Journal of the American Chemical Society 129, 13713-13722.

Mann, S. (2001). Biomineralization : principles and concepts in bioinorganic materials chemistry. Oxford: Oxford University Press.

Marchin, K. L. & Berrie, C. L. (2003). Langmuir 19, 9883-9888.

Martin, R. B., Burr, D. B. & Sharkey, N. A. (1998). Skeletal tissue mechanics. New York: Springer.

Mason, P. E., Cruickshank, J. M., Neilson, G. W. & Buchanan, P. (2003). Physical Chemistry Chemical Physics 5, 4686-4690.

Mayo, S. L., Olafson, B. D. & Goddard, W. A. (1990). Journal of Physical Chemistry 94, 8897-8909.

Meis, C., Gale, J. D., Boyer, L., Carpena, J. & Gosset, D. (2000). Journal of Physical Chemistry A 104, 5380-5387.

Minton, A. P. (1999). Biophys J 76, 176-187.

Mkhonto, D. & de Leeuw, N. H. (2002). Journal of Materials Chemistry 12, 2633-2642.

Mkhonto, D., Ngoepe, P. E., Cooper, T. G. & de Leeuw, N. H. (2006). Physics and Chemistry of Minerals 33, 314-331.

Mostafa, N. Y. & Brown, P. W. (2006). Journal of Physics and Chemistry of Solids 68, 431-437.

Murugan, R. & Ramakrishna, S. (2005). Crystal Growth & Design 5, 111-112.

Nagarajan, V., Kamitori, S. & Okuyama, K. (1998). Journal of Biochemistry 124, 1117-1123.

Nair, N. N., Schreiner, E. & Marx, D. (2006). Journal of the American Chemical Society 128, 13815-13826.

Nakayama, S., Kageyama, T., Aono, H. & Sadaoka, Y. (1995). Journal of Materials Chemistry 5, 1801-1805.

Narasaraju, T. S. B. & Phebe, D. E. (1996). J Mater Sci 31, 1-21.

Nose, S. (1984). Journal of Chemical Physics 81, 511-519.

Okuyama, K., Hongo, C., Fukushima, R., Wu, G. G., Narita, H., Noguchi, K., Tanaka, Y. & Nishino, N. (2004). Biopolymers 76, 367-377.

Ordejon, P., Artacho, E. & Soler, J. M. (1996). Phys Rev B 53, 10441-10444.

Ozolins, V. & Korling, M. (1993). Physical Review B 48, 18304-18307.

Pan, H. H., Tao, J. H., Xu, X. R. & Tang, R. K. (2007). Langmuir 23, 8972-8981.

Park, C. Y., Fenter, P., Zhang, Z., Cheng, L. W. & Sturchio, N. C. (2004). American Mineralogist 89, 1647-1654.

Parr, R. G. & Yang, a. W. (1989). Density Functional Theory of Atoms and Molecules. New York: Oxford University Press.

Perdew, J. P. (1986). Physical Review B 33, 8822-8824.

Perdew, J. P., Burke, K. & Ernzerhof, M. (1996). Phys Rev Lett 77, 3865-3868.

Perdew, J. P., Chevary, J. A., Vosko, S. H., Jackson, K. A., Pederson, M. R., Singh, D. J. & Fiolhais, C. (1993). Physical Review B 48, 4978-4978.

Perdew, J. P., McMullen, E. R. & Zunger, A. (1981). Physical Review A 23, 2785-2789.

Perdew, J. P. & Wang, Y. (1992). Physical Review B 45, 13244-13249.

Perdew, J. P. & Yue, W. (1986). Physical Review B 33, 8800-8802.

Pisani C., D. R., Roetti C., (1996). Quantum Mechanical Calculations of the Properties of Crystalline Materials. Springer.

Posner, A. S., Perloff, A. & Diorio, A. F. (1958). Acta Crystallographica 11, 308-309.

Prockop, D. J. & Kivirikko, K. I. (1995). Annu Rev Biochem 64, 403-434.

Prokopowicz, M., Lukasiak, J., Banecki, B. & Przyjazny, A. (2005). Biomacromolecules 6, 39-45.

Qian, J. J. & Bhatnagar, R. S. (1996). Journal of Biomedical Materials Research 31, 545-554.

Qiu, X. Y., Hong, Z. K., Hu, J. L., Chen, L., Chen, X. S. & Jing, X. B. (2005). Biomacromolecules 6, 1193-1199.

Quirk, R. A., France, R. M., Shakesheff, K. M. & Howdle, S. M. (2004). Current Opinion in Solid State & Materials Science 8, 313-321.

Rautaray, D., Mandal, S. & Sastry, M. (2005). Langmuir 21, 5185-5191.

Ravichandran, S., Madura, J. D. & Talbot, J. (2001). Journal of Physical Chemistry B 105, 3610-3613.

Rehman, I. & Bonfield, W. (1997). Journal of Materials Science-Materials in Medicine 8, 1-4.

Remko, M. & Rode, B. M. (2006). Journal of Physical Chemistry A 110, 1960-1967.

Rimola, A., Corno, M., Zicovich-Wilson, C. & Ugliengo, P. (2008). Journal of the American Chemical Society 130, 16181–16183.

Robinson, J., Cukrowski, I. & Marques, H. M. (2006). Journal of Molecular Structure 825, 134-142.

Sanders, D. M., Person, W. B. & Hench, L. L. (1972). Applied Spectroscopy 26, 530-&.

Sanger, A. T. & Kuhs, W. F. (1992). Zeitschrift Fur Kristallographie 199, 123-148.

Sarikaya, M. & Aksay, I. A. (1995). Biomimetics : design and processing of materials. Woodbury, N.Y.: AIP Press.

Sato, K., Kogure, T., Iwai, H. & Tanaka, J. (2002). Journal of the American Ceramic Society 85, 3054-3058.

Schroder, K. P., Sauer, J., Leslie, M., Catlow, C. R. A. & Thomas, J. M. (1992). Chemical Physics Letters 188, 320-325.

Shaw, W. J., Long, J. R., Dindot, J. L., Campbell, A. A., Stayton, P. S. & Drobny, G. P. (2000). Journal of the American Chemical Society 122, 1709-1716.

Shen, J. W., Wu, T., Wang, Q. & Pan, H. H. (2008). Biomaterials 29, 513-532.

Simon, P., Schwarz, U. & Kniep, R. (2005). Journal of Materials Chemistry 15, 4992-4996.

Simon, P., Zahn, D., Lichte, H. & Kniep, M. (2006). Angewandte Chemie-International Edition 45, 1911-1915.

Singh, D. J. & Ashkenazi, J. (1992). Physical Review B 46, 11570-11577.

Slater, J. C. (1930). Physical Review 35, 210.

Smith, C. E. & Nanci, A. (1996). Anat Rec 245, 186-207.

Smith, W. & Forester, T. R. (1996). Journal of Molecular Graphics 14, 136-141.

Soler, J. M., Artacho, E., Gale, J. D., Garcia, A., Junquera, J., Ordejon, P. & Sanchez-Portal, D. (2002). Journal of Physics-Condensed Matter 14, 2745-2779.

Stevens, M. M. & George, J. H. (2005). Science 310, 1135-1138.

Stork, L., Muller, P., Dronskowski, R. & Ortlepp, J. R. (2005). Zeitschrift Fur Kristallographie 220, 201-205.

Swope, W. C., Andersen, H. C., Berens, P. H. & Wilson, K. R. (1982). Journal of Chemical Physics 76, 637-649.

Tanaka, H., Miyajima, K., Nakagaki, M. & Shimabayashi, S. (1989). Chemical & Pharmaceutical Bulletin 37, 2897-2901.

Tang, E., Di Tommaso, D. & de Leeuw, N. H. (2009a). Journal of Chemical Physics In press.

Tang, E., Di Tommaso, D. & de Leeuw, N. H. (2009b). Journal of Chemical Physics 130.

Tarasevich, B. J., Howard, C. J., Larson, J. L., Snead, M. L., Simmer, J. P., Paine, M. & Shaw, W. J. (2007). Journal of Crystal Growth 304, 407-415.

Tasker, P. W. (1979). Philosophical Magazine a-Physics of Condensed Matter Structure Defects and Mechanical Properties 39, 119-136.

Tirrell, M., Kokkoli, E. & Biesalski, M. (2002). Surf Sci 500, 61-83.

Tlatlik, H., Simon, P., Kawska, A., Zahn, D. & Kniep, R. (2006). Angewandte Chemie-International Edition 45, 1905-1910.

Todorov, I. T., Smith, W., Trachenko, K. & Dove, M. T. (2006). Journal of Materials Chemistry 16, 1911-1918.

Tolchard, J. R., Slater, P. R. & Islam, M. S. (2007). Advanced Functional Materials 17, 2564-2571.

Troullier, N. & Martins, J. L. (1991). Phys Rev B 43, 8861-8869.

Tsuda, H. & Arends, J. (1994). Journal of Dental Research 73, 1703-1710.

Vanderby, R. (2003). J Biomech 36, 1523-1527.

Verlet, L. (1967). Physical Review 159, 98-&.

Vitagliano, L., Berisio, R., Mazzarella, L. & Zagari, A. (2001). Biopolymers 58, 459-464.

Vocadlo, L., de Wijs, G. A., Kresse, G., Gillan, M. J. & Price, G. D. (1997). Faraday Discuss. 106, 205-217.

Vosko, S. H., Wilk, L. & Nusair, M. (1980). Canadian Journal of Physics 58, 1200-1211.

Walser, R., Hunenberger, P. H. & van Gunsteren, W. F. (2001). Proteins-Structure Function and Genetics 43, 509-519.

Wang, A. Y., Mo, X., Chen, C. S. & Yu, S. M. (2005). Journal of the American Chemical Society 127, 4130-4131.

Watson, G. W., Kelsey, E. T., deLeeuw, N. H., Harris, D. J. & Parker, S. C. (1996). J Chem Soc Faraday T 92, 433-438.

Wen, H. B., Fincham, A. G. & Moradian-Oldak, J. (2001). Matrix Biol 20, 387-395.

Wierzbicki, A. & Cheung, H. S. (2000). J Mol Struc-Theochem 529, 73-82.

Xia, N. & Castner, D. G. (2003). Journal of Biomedical Materials Research Part A 67A, 179-190.

Yu, D., Armstrong, D. A. & Rauk, A. (1992). Can J Chem 70, 1762-1772.

Zahn, D. & Hochrein, O. (2003). Physical Chemistry Chemical Physics 5, 4004-4007.

Zahn, D., Hochrein, O., Kawska, A., Brickmann, J. & Kniep, R. (2007). J Mater Sci 42, 8966-8973.

Zapantal.R (1965). Nature 206, 403-&.